

SEARCH REQUEST FORM

Scientific and Technical Information Center

Requester's Full Name: _____ Examiner #: _____ Date: _____
 An Unit: _____ Phone Number 30 _____ Serial Number: _____
 Mail Box and Bldg/Room Location: _____ Results Format Preferred (circle): PAPER DISK E-MAIL

If more than one search is submitted, please prioritize searches in order of need.

Please provide a detailed statement of the search topic, and describe as specifically as possible the subject matter to be searched. Include the elected species or structures, keywords, synonyms, acronyms, and registry numbers, and combine with the concept or utility of the invention. Define any terms that may have a special meaning. Give examples or relevant citations, authors, etc. if known. Please attach a copy of the cover sheet, pertinent claims, and abstract.

Title of Invention: _____

Inventors (please provide full names): _____

Earliest Priority Filing Date: _____

**For Sequence Searches Only* Please include all pertinent information (parent, child, divisional, or issued patent numbers) along with the appropriate serial number.*

STAFF USE ONLY

Searcher: P. Schreber

Searcher Phone #: 308-4292

Searcher Location: CM1 12C14

Date Searcher Provided: _____

Date Completed: 12/10

Searcher Prep & Review Time: 25

Client Prep Time: _____

Printing Time: 7

PT: _____

Type of Search

NA Sequence (#) 7

AA Sequence (#) 1

Structure (#) _____

Bibliographic _____

Litigation _____

Fulltext: _____

Patent Family _____

Other: _____

Vendors and cost where applicable

STN: _____

Dialog: _____

Questel Orbit: _____

Dr. Link: _____

John's News: _____

Sequence Systems: CompuGen

WVA/Internet: _____

Other Specialty: _____

THIS PAGE BLANK (USPTO)

XX

Protein encoded by
Protein encoded by
Protein encoded by
Protein encoded by
Protein encoded by
Protein encoded by
Arabidopsis thaliana
Arabidopsis thaliana
C. roseus DAT protob
Aromatic acyl tran
Polypeptide encoded
Tobacco Ant23 genom
Aromatic acyl tran
Arabidopsis thaliana
Aromatic acyl tran
Arabidopsis thaliana
Honey dew melon al
Aromatic acyl tran
Arabidopsis thaliana
Arabidopsis thaliana
Arabidopsis thaliana
Arabidopsis thaliana
Arabidopsis thaliana
Arabidopsis thaliana
C. roseus acyl tra
Arabidopsis thaliana
Arabidopsis thaliana
Banana alcohol acy
Arabidopsis thaliana
Arabidopsis thaliana
Arabidopsis thaliana
Arabidopsis thaliana
Arabidopsis thaliana

PT DNA coding for aromatic acyl transferase - for transforming plants
PT which produce anthocyanin pigments and thus altering colour tone,
e.g. of flowers
XX
PS Claim 2: Page 81: 94pp: Japanese.
XX
CC Vectors containing DNA fragments encoding proteins of plant origin
CC with aromatic acyl transferase activity may be used to transform
CC plants which produce anthocyanin pigments. The aromatic acyl
CC transferase acylates the pigments in the flower resulting in colour
CC tone changes and allowing new colourations to be produced. Six
CC specific DNA sequences encoding aromatic acyl transferase from
CC different plants are described in AAT37308-F37313. This peptide
CC fragment was isolated from three of the six clones and was used to
CC synthesise a degenerate primer (AAT37314) which was then used to
CC identify other aromatic acyl transferase encoding clones.
SQ Sequence 6 AA:

Query Match 100.0%; Score 40; DB 17; Length 6;
Best Local Similarity 100.0%; Pred. No. 4.3e+05;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DFGWGK 6
| | | | |
Db 1 dfgwgk 6

RESULT 2
AAE00248
ID AAE00248 standard; peptide; 7 AA.
XX
AC AAE00248;
XX
DT 13-JUN-2001 (first entry)
XX
DE Peptide fragment #3 to construct degenerate reverse PCR primer AT-REV1.
XX
KM Transacylase: taxol; pacilitaxel biosynthesis; taxoid;
KM transgenic organism.
XX
OS Unidentified.
XX
PN WO200123586-A2.
XX
PD 05-APR-2001.
XX
PF 29-SEP-2000; 2000WO-US27006.
XX
PR 30-SEP-1999; 99US-0411145.
PR 07-DEC-1999; 99US-0457046.
XX
PA (UNITM) UNITV WASHINGTON STATE RES FOUND.
XX
XX Croteau RB, Walker KD, Schoendorf A, Wildung MR;
PI WPI: 2001-245004/25.
XX
DR New transacylase enzymes, useful for the high yield production of
PT Taxol(TM), related taxoids and useful intermediates in the in the
PT pacilitaxel biosynthetic pathway -
XX
XX
PS Disclosure: Page 40; 162pp: English.
XX
CC The present sequence is a peptide fragment used for constructing
CC degenerate reverse PCR primer AT-REV1. The primers are useful for
CC generating probes which are useful for the identification of
CC (nucleic acid sequences encoding) transacylases. The probes
CC isolated from the Taxus genus are useful for the synthetic production
CC of Taxol(TM) and related taxoids, as well as intermediates in the
CC pacilitaxel biosynthetic pathway. They can also be used for the creation
CC of transgenic organisms that either produce the transacylases for

CC subsequent in vitro use, or produce the transacylases in vivo. The
CC (nucleic acids encoding) transacylases are also useful for creating
CC specific binding agents that recognise the corresponding transacylases.
CC Binding agents include (fragments of) antibodies or any other agent
CC capable of specifically binding to the groups on the proteins.
XX
SQ Sequence 7 AA:

Query Match 100.0%; Score 40; DB 22; Length 7;
Best Local Similarity 100.0%; Pred. No. 4.3e+05;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DFGWGK 6
| | | | |
Db 1 dfgwgk 6

RESULT 3
AAG07709
ID AAG07709 standard; Protein; 287 AA.
XX
AC AAG07709;
XX
DT 17-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 4964.
XX
KM Protein identification: signal transduction pathway; metabolic pathway;
KM hybridisation assay; genetic mapping; gene expression control; promoter;
KM termination sequence.
XX
XX Arabidopsis thaliana.
OS EPI033405-A2.
XX
PN EPI033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-0301439.
XX
PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0123788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0128845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 28-APR-1999; 99US-0130891.
PR 30-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 04-MAY-1999; 99US-0132407.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 07-MAY-1999; 99US-0132487.
PR 11-MAY-1999; 99US-0132863.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.

PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 14-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139992.
PR 18-JUN-1999; 99US-0139934.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145228.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.

PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 25-AUG-1999; 99US-0149930.
PR 26-AUG-1999; 99US-0150566.
PR 27-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 30-AUG-1999; 99US-0151080.
PR 31-AUG-1999; 99US-0151303.
PR 01-SEP-1999; 99US-0151438.
PR 07-SEP-1999; 99US-0151930.
PR 10-SEP-1999; 99US-0153363.
PR 13-SEP-1999; 99US-0153070.
PR 15-SEP-1999; 99US-0153758.
PR 16-SEP-1999; 99US-0154018.
PR 20-SEP-1999; 99US-0154039.
PR 22-SEP-1999; 99US-0154779.
PR 23-SEP-1999; 99US-0155139.
PR 24-SEP-1999; 99US-0155486.
PR 28-SEP-1999; 99US-0155659.
PR 29-SEP-1999; 99US-0156458.
PR 04-OCT-1999; 99US-0156596.
PR 05-OCT-1999; 99US-0157117.
PR 06-OCT-1999; 99US-0157753.
PR 07-OCT-1999; 99US-0157865.
PR 08-OCT-1999; 99US-0158029.
PR 12-OCT-1999; 99US-0158232.
PR 13-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 14-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159638.
PR 21-OCT-1999; 99US-0159741.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.

```
PR 29-OCT-1999; 99US-0162142.
Query Match 100.0%; Score 40; DB 21; Length 287;
Best Local Similarity 100.0%; Pred. No. 26;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 DEFGMK 6
    |||||
Db 229 dfgwqk 234

RESULT 4
AAG49296
ID AAG49296 standard; Protein: 287 AA.
XX
XX AAG49296;
AC
AC 18-OCT-2000 (first entry)
DE
DE Arabidopsis thaliana protein fragment SEQ ID NO: 62350.
KW
KW Protein identification; signal transduction pathway; metabolic pathway;
KW hydridisation assay; genetic mapping; gene expression control; promoter;
KM termination sequence.
XX
XX Arabidopsis thaliana.
OS
PN EPI033405-A2.
XX
PD 06-SEP-2000.
XX
XX 25-FEB-2000; 2000EP-0301439.
XX
XX 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0128845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 28-APR-1999; 99US-0130891.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132862.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
```

```
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139753.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144088.
PR 16-JUL-1999; 99US-0144088.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 28-JUL-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
```

```

db      229 dfgwqk 234
|||||
RESULT 5
ANG07708
ID     AGC07708 standard; Protein: 293 AA.
XX
AC     AAC07708;
XX
DT     17-OCT-2000 (first entry)
DE     Arabidopsis thaliana protein fragment SEQ ID NO: 4963.
XX
KW     Protein identification; signal transduction pathway; metabolic pathway;
KW     hybridisation assay; genetic mapping; gene expression control; promoter;
XX     termination sequence.
XX
OS     Arabidopsis thaliana.
XX
PM     EP1033405-A2.
XX
PD     06-SEP-2000.
XX
PF     25-FEB-2000; 2000EP-0301439.
XX
PR     25-FEB-1999; 99US-0121825.
PR     05-MAR-1999; 99US-0123180.
PR     09-MAR-1999; 99US-0123548.
PR     23-MAR-1999; 99US-0125788.
PR     25-MAR-1999; 99US-0126264.
PR     29-MAR-1999; 99US-0126785.
PR     01-APR-1999; 99US-0127462.
PR     06-APR-1999; 99US-0128234.
PR     08-APR-1999; 99US-0128714.
PR     16-APR-1999; 99US-0129845.
PR     19-APR-1999; 99US-0130077.
PR     21-APR-1999; 99US-0130449.
PR     23-APR-1999; 99US-0130510.
PR     28-APR-1999; 99US-0130891.
PR     30-APR-1999; 99US-0132068.
PR     04-MAY-1999; 99US-0132407.
PR     05-MAY-1999; 99US-0132484.
PR     06-MAY-1999; 99US-0132485.
PR     06-MAY-1999; 99US-0132486.
PR     07-MAY-1999; 99US-0132487.
PR     11-MAY-1999; 99US-0134256.
PR     14-MAY-1999; 99US-0134218.
PR     14-MAY-1999; 99US-0134219.
PR     14-MAY-1999; 99US-0134221.
PR     18-MAY-1999; 99US-0134370.
PR     19-MAY-1999; 99US-0134768.
PR     20-MAY-1999; 99US-0134941.
PR     21-MAY-1999; 99US-0135124.
PR     24-MAY-1999; 99US-0135353.
PR     25-MAY-1999; 99US-0135629.
PR     27-MAY-1999; 99US-0136021.
PR     28-MAY-1999; 99US-0136392.
PR     01-JUN-1999; 99US-0136782.
PR     03-JUN-1999; 99US-0137222.
PR     04-JUN-1999; 99US-0137528.
PR     07-JUN-1999; 99US-0137502.
PR     08-JUN-1999; 99US-0137724.
PR     10-JUN-1999; 99US-0138094.
PR     20-JUN-1999; 99US-0138180.

```

PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.

XX AC AAG49295;
XX 18-OCT-2000 (first entry)
XX Arabidopsis thaliana protein fragment SEQ ID NO: 62349.
DE Protein Identification: signal transduction pathway; metabolic pathway;
XX hybridisation assay; genetic mapping; gene expression control; promoter;
XX termination sequence.
OS Arabidopsis thaliana.
XX
XX EP1033405-A2.
PD 06-SEP-2000.
XX
XX 25-FEB-2000; 2000EP-0301439.
XX
XX 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.

PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144632.
PR 21-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 03-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148584.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.


```

PF 29-SEP-2000: 2000MO-US27006.
XX
XX 30-SEP-1999: 99US-0411145.
PR 07-DEC-1999: 99US-0457046.
XX
XX (UNITM ) UNIV WASHINGTON STATE RES FOUND.
PI Croteau RB, Walker KD, Schoendorf A, Wildung MR:
DR WPI: 2001-245004/25.
XX
XX New transacylase enzymes, useful for the high yield production of
PT Taxol(TM), related taxoids and useful intermediates in the in the
PT pacilitaxel biosynthetic pathway
XX
XX Claim 1: Page 102-103; 162pp: English.
PS
XX
XX The present sequence is the amino acid sequence of probe 5. Probe 5 is
CC derived from AT-FOR4 and AT-REV1 primers and is used for screening
CC Taxus cuspidata TAX5 full length cDNA clone. The probes are useful for
CC the identification of (nucleic acid sequences encoding) transacylases.
CC The probes isolated from the Taxus genus are useful for the
CC synthetic production of Taxol(TM) and related taxoids, as well as
CC intermediates in the pacilitaxel biosynthetic pathway. They can also be
CC used for the creation of transgenic organisms that either produce the
CC transacylases for subsequent in vitro use, or produce the transacylases
CC in vivo. The (nucleic acids encoding) transacylases are also useful for
CC creating specific binding agents that recognise the corresponding
CC transacylases. Binding agents include (fragments of) antibodies or
CC any other agent capable of specifically binding to the groups on
CC the proteins.
CC Note: The specification states that this sequence is encoded by the
CC sequence shown in AAD03337. However, this does not appear to be the case.
XX
XX Sequence 302 AA:
SQ

```

```

Query Match          100.0%; Score 40; DB 22; Length 302;
Best Local Similarity 100.0%; Pred. No. 27;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 1 DFGMGK 6
   |||||
DB 297 dfwgk 302

```

```

RESULT 9
AAE00234
ID AAE00234 standard; Protein: 302 AA.
XX
XX AAE00234:
AC
XX
XX 13-JUN-2001 (first entry)
DT
XX
XX Protein encoded by Probe 8.
DE
XX
XX Transacylase; taxol; pacilitaxel biosynthesis; taxoid; probe;
KW transgenic organism.
XX
XX Taxus cuspidata.
OS
XX
XX WO200123586-A2.
PN
XX
XX 05-APR-2001.
PD
XX
XX 29-SEP-2000: 2000MO-US27006.
PF
XX
XX 30-SEP-1999: 99US-0411145.
PR 07-DEC-1999: 99US-0457046.
XX
XX (UNITM ) UNIV WASHINGTON STATE RES FOUND.
PA
XX
XX Croteau RB, Walker KD, Schoendorf A, Wildung MR:
PI

```

```

XX
XX WPI: 2001-245004/25.
DR N-PSDB; AAD03340.
XX
XX New transacylase enzymes, useful for the high yield production of
PT Taxol(TM), related taxoids and useful intermediates in the in the
PT pacilitaxel biosynthetic pathway
XX
XX Claim 1: Page 107-108; 162pp: English.
PS
XX
XX The present sequence is the amino acid sequence of probe 8.
CC Probe 8 is derived from AT-FOR3 and AT-REV1 primers and is
CC used for screening Taxus cuspidata transacylase cDNA. The
CC probes are useful for the identification of (nucleic acid sequences
CC encoding) transacylases. The probes isolated from the Taxus genus
CC are useful for the synthetic production of Taxol(TM) and related taxoids,
CC as well as intermediates in the pacilitaxel biosynthetic pathway. They
CC can also be used for the creation of transgenic organisms that either
CC produce the transacylases for subsequent in vitro use, or produce the
CC transacylases in vivo. The (nucleic acids encoding) transacylases are
CC also useful for creating specific binding agents that recognise the
CC corresponding transacylases. Binding agents include (fragments of)
CC antibodies or any other agent capable of specifically binding to the
CC groups on the proteins.
XX
XX Sequence 302 AA:
SQ

```

```

Query Match          100.0%; Score 40; DB 22; Length 302;
Best Local Similarity 100.0%; Pred. No. 27;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 1 DFGMGK 6
   |||||
DB 297 dfwgk 302

```

```

RESULT 10
AAE00235
ID AAE00235 standard; Protein: 302 AA.
XX
XX AAE00235:
AC
XX
XX 13-JUN-2001 (first entry)
DT
XX
XX Protein encoded by Probe 9.
DE
XX
XX Transacylase; taxol; pacilitaxel biosynthesis; taxoid; probe;
KW transgenic organism; TAX9.
XX
XX Taxus cuspidata.
OS
XX
XX FH Key Location/Qualifiers
FT Misc-difference 119 /note="Encoded by TGG"
FT Misc-difference 164 /note="Encoded by ATG"
FT Misc-difference 187 /note="Encoded by ATG"
FT Misc-difference 187 /note="Encoded by TGC"
XX
XX WO200123586-A2.
PN
XX
XX 05-APR-2001.
PD
XX
XX 29-SEP-2000: 2000MO-US27006.
PF
XX
XX 30-SEP-1999: 99US-0411145.
PR 07-DEC-1999: 99US-0457046.
XX
XX (UNITM ) UNIV WASHINGTON STATE RES FOUND.
PA
XX
XX Croteau RB, Walker KD, Schoendorf A, Wildung MR:
PI

```

DR WPI: 2001-245004/25.
 DR N-PSDB: AAD03341.
 PT New transacylase enzymes, useful for the high yield production of
 PT Taxol(TM), related taxoids and useful intermediates in the in the
 PT pacitaxel biosynthetic pathway -
 PS Claim 1: Page 108-109; 162pp; English.
 XX
 XX The present sequence is the amino acid sequence of probe 9.
 CC Probe 9 is derived from AT-FOR2 and AT-REVL primers and is
 CC used for screening Taxus cuspidata TAX9 full length cDNA clone. The
 CC probes are useful for the identification of (nucleic acid sequences
 CC encoding) transacylases. The probes isolated from the Taxus genus
 CC are useful for the synthetic production of Taxol(TM) and related taxoids,
 CC as well as intermediates in the pacitaxel biosynthetic pathway. They
 CC can also be used for the creation of transgenic organisms that either
 CC produce the transacylases for subsequent in vitro use, or produce the
 CC transacylases in vivo. The (nucleic acids encoding) transacylases are
 CC also useful for creating specific binding agents that recognise the
 CC corresponding transacylases. Binding agents include (fragments of)
 CC antibodies or any other agent capable of specifically binding to the
 CC groups on the proteins.
 CC
 SQ Sequence 302 AA;

Query Match 100.0%; Score 40; DB 22; Length 302;
 Best Local Similarity 100.0%; Pred. No. 27;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 DPGWCK 6
 |||||
 Db 297 dfgwck 302

RESULT 11

AAE00238
 ID AAE00238 standard; Protein: 302 AA.

AC AAE00238;

DT 13-JUN-2001 (first entry)

DE Protein encoded by Probe 12.

XX Transacylase; taxol; pacitaxel biosynthesis; taxoid; probe;

KW Transgenic organism; TAX12.

XX Taxus cuspidata.

OS WO200123586-A2.

XX 05-APR-2001.

PD 29-SEP-2000; 2000WO-US27006.

XX 30-SEP-1999; 99US-0411145.

PR 07-DEC-1999; 99US-0457046.

XX (UNIW) UNIV WASHINGTON STATE RES FOUND.

PA Croteau RB, Walker KD, Schoendorf A, Wildung MR;

XX WPI: 2001-245004/25.

DR N-PSDB: AAD03344.

XX New transacylase enzymes, useful for the high yield production of
 PT Taxol(TM), related taxoids and useful intermediates in the in the
 PT pacitaxel biosynthetic pathway -
 PS Claim 1; Page 113; 162pp; English.

CC The present sequence is the amino acid sequence of probe 12. Probe 12
 CC is derived from AT-FOR3 and AT-REVL primers and is used for screening
 CC Taxus cuspidata TAX12 full length cDNA clone. The probes are useful for
 CC the identification of (nucleic acid sequences encoding) transacylases.
 CC The probes isolated from the Taxus genus are useful for the
 CC synthetic production of Taxol(TM) and related taxoids, as well as
 CC intermediates in the pacitaxel biosynthetic pathway. They can also
 CC be used for the creation of transgenic organisms that either
 CC produce the transacylases for subsequent in vitro use, or produce the
 CC transacylases in vivo. The (nucleic acids encoding) transacylases are
 CC also useful for creating specific binding agents that recognise the
 CC corresponding transacylases. Binding agents include (fragments of)
 CC antibodies or any other agent capable of specifically binding to the
 CC groups on the proteins.
 CC
 SQ Sequence 302 AA;

Query Match 100.0%; Score 40; DB 22; Length 302;
 Best Local Similarity 100.0%; Pred. No. 27;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 DPGWCK 6
 |||||
 Db 297 dfgwck 302

RESULT 12

AAE00232
 ID AAE00232 standard; Protein: 303 AA.

AC AAE00232;

DT 13-JUN-2001 (first entry)

DE Protein encoded by Probe 6.

XX Transacylase; taxol; pacitaxel biosynthesis; transgenic organism; TAX6;
 KW (10-deacetyl)baccalin III-0-O-acetyltransferase); taxoid; probe.

XX Taxus cuspidata.

OS WO200123586-A2.

XX 05-APR-2001.

PD 29-SEP-2000; 2000WO-US27006.

XX 30-SEP-1999; 99US-0411145.

PR 07-DEC-1999; 99US-0457046.

XX (UNIW) UNIV WASHINGTON STATE RES FOUND.

PA Croteau RB, Walker KD, Schoendorf A, Wildung MR;

XX WPI: 2001-245004/25.

DR N-PSDB: AAD03338.

XX New transacylase enzymes, useful for the high yield production of
 PT Taxol(TM), related taxoids and useful intermediates in the in the
 PT pacitaxel biosynthetic pathway -
 PS Claim 1; Page 104-105; 162pp; English.
 XX
 XX The present sequence is the amino acid sequence of probe 6.
 CC Probe 6 is derived from AT-FOR2 and AT-REVL primers and is
 CC used for screening Taxus cuspidata full length cDNA clone
 CC TAX6 (10-deacetyl)baccalin III-0-O-acetyl transferase). The
 CC probes are useful for the identification of (nucleic acid sequences
 CC encoding) transacylases. The probes isolated from the Taxus genus
 CC are useful for the synthetic production of Taxol(TM) and related taxoids,
 CC as well as intermediates in the pacitaxel biosynthetic pathway. They
 CC can also be used for the creation of transgenic organisms that either

CC produce the transacylases for subsequent in vitro use, or produce the
CC transacylases in vivo. The (nucleic acids encoding) transacylases are
CC also useful for creating specific binding agents that recognise the
CC corresponding transacylases. Binding agents include (fragments of)
CC antibodies or any other agent capable of specifically binding to the
CC groups on the proteins.

XX Sequence 303 AA:

Query Match 100.0%; Score 40; DB 22; Length 303;

Best Local Similarity 100.0%; Pred. No. 27;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 DFGWGK 6
|||||
Db 298 dfgwsk 303

RESULT 13

AAE00236
ID AAE00236 standard; Protein: 303 AA.

XX
AC AAE00236;

XX 13-JUN-2001 (first entry)

XX Protein encoded by Probe 10.

DE Transacylase; taxol; paciltaxel biosynthesis; taxoid; probe;

XX Transgenic organism; TAX10.

XX Taxus cuspidata.

XX WO200123586-A2.

XX 05-APR-2001.

XX 29-SEP-2000; 2000WO-US27006.

XX 30-SEP-1999; 99US-0411145.

XX 07-DEC-1999; 99US-0457046.

PA (UNITW) UNIV WASHINGTON STATE RES FOUND.

PI Croceanu RB, Walker KD, Schoendorf A, Wildung MR;

DR WPI: 2001-245004/25.

DR N-PSDB: AAD03342.

XX New transacylase enzymes, useful for the high yield production of
XX Taxol(TM), related taxoids and useful intermediates in the in the
XX paciltaxel biosynthetic pathway -

PS Claim 1; Page 110-111; 162pp; English.

XX The present sequence is the amino acid sequence of probe 10.

XX Probe 10 is derived from AT-FOR4 and AT-REV1 primers is used
XX for screening Taxus cuspidata TAX10 full length cDNA clone. The
XX probes are useful for the identification of (nucleic acid sequences
XX encoding) transacylases. The probes isolated from the Taxus genus
XX are useful for the synthetic production of Taxol(TM) and related taxoids,
XX as well as intermediates in the paciltaxel biosynthetic pathway. They
XX can also be used for the creation of transgenic organisms that either
XX produce the transacylases for subsequent in vitro use, or produce the
XX transacylases in vivo. The (nucleic acids encoding) transacylases are
XX also useful for creating specific binding agents that recognise the
XX corresponding transacylases. Binding agents include (fragments of)
XX antibodies or any other agent capable of specifically binding to the
XX groups on the proteins.

SO Sequence 303 AA:

Query Match 100.0%; Score 40; DB 22; Length 303;
Best Local Similarity 100.0%; Pred. No. 27;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 DFGWGK 6
|||||
Db 298 dfgwsk 303

RESULT 14

AAE00227
ID AAE00227 standard; Protein: 306 AA.

XX
AC AAE00227;

XX 13-JUN-2001 (first entry)

XX Protein encoded by Probe 1.

DE TAX1; taxadien-5 alpha-ol acetyl transferase; transacylase; taxol;

XX TAX2; taxadien-2 alpha-O-benzoyl transferase; paciltaxel biosynthesis;

XX taxoid; Transgenic organism; probe.

XX Taxus cuspidata.

XX WO200123586-A2.

XX 05-APR-2001.

XX 29-SEP-2000; 2000WO-US27006.

XX 30-SEP-1999; 99US-0411145.

XX 07-DEC-1999; 99US-0457046.

PA (UNITW) UNIV WASHINGTON STATE RES FOUND.

PI Croceanu RB, Walker KD, Schoendorf A, Wildung MR;

DR WPI: 2001-245004/25.

DR N-PSDB: AAD03333.

XX The present sequence is the amino acid sequence of probe 1. Probe 1 is
XX derived from AT-FOR1 and AT-REV1 primers and is used for screening
XX Taxus cuspidata full length cDNA clones TAX1 or TAX01 (taxadien-5 alpha-
XX ol acetyl transferase) and TAX2 or TAX02 (taxadien-2 alpha-O-benzoyl
XX transferase). The probes are useful for the identification of (nucleic
XX acid sequences encoding) transacylases. The probes isolated from
XX the Taxus genus are useful for the synthetic production of Taxol(TM) and
XX related taxoids, as well as intermediates in the paciltaxel biosynthetic
XX pathway. They can also be used for the creation of transgenic organisms
XX that either produce the transacylases for subsequent in vitro use, or
XX produce the transacylases in vivo. The (nucleic acids encoding)
XX transacylases are also useful for creating specific binding agents that
XX recognise the corresponding transacylases. Binding agents include
XX (fragments of) antibodies or any other agent capable of specifically
XX binding to the groups on the proteins.

SO Sequence 306 AA:

Query Match 100.0%; Score 40; DB 22; Length 306;
Best Local Similarity 100.0%; Pred. No. 27;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 DFGWGK 6
|||||

Db 301 dfgwsk 306

RESULT 15

AAE00228
ID AAE00228 standard; Protein; 306 AA.

XX AAE00228;

XX 13-JUN-2001 (first entry)

XX Protein encoded by Probe 2.

XX TAX2; taxadien-2 alpha-O-benzoyl transferase; pacilitaxel biosynthesis;
KM taxold; transgenic organism; probe; transacylase; taxol.

XX Taxus cuspidata.

XX WO200123586-A2.

XX 05-APR-2001.

XX 29-SEP-2000; 2000WO-US27006.

XX 30-SEP-1999; 99US-0411145.
PR 07-DEC-1999; 99US-0457046.

XX (UNIW) UNIV WASHINGTON STATE RES FOUND.

XX Croteau RB, Walker KD, Schoendorf A, Wildung MR;

XX WPI; 2001-245004/25.

XX N-PSDB; AAD03334.

XX New transacylase enzymes, useful for the high yield production of
PT Taxol(TM), related taxoids and useful intermediates in the in the
PT pacilitaxel biosynthetic pathway -

PS Claim 1; Page 98-99; 162pp; English.

XX The present sequence is the amino acid sequence of probe 2. Probe 2
CC is used for screening Taxus cuspidata transacylase cDNA. Probe 2
CC is derived from AT-FOR1 and AT-REV1 primers but it is not useful
CC for screening TAX2 or TAX02 (taxadien-2 alpha-O-benzoyl
CC transferase) but has the sequence similar to TAX2 gene. The
CC probes are useful for the identification of (nucleic acid sequences
CC encoding) transacylases. The probes isolated from the Taxus genus
CC are useful for the synthetic production of Taxol(TM) and related taxoids,
CC as well as intermediates in the pacilitaxel biosynthetic pathway. They
CC can also be used for the creation of transgenic organisms that either
CC produce the transacylases for subsequent in vitro use, or produce the
CC transacylases in vivo. The (nucleic acids encoding) transacylases are
CC also useful for creating specific binding agents that recognise the
CC corresponding transacylases. Binding agents include (fragments of)
CC antibodies or any other agent capable of specifically binding to the
CC groups on the proteins.

XX Sequence 306 AA;

Query March 100.0%; Score 40; DB 22; Length 306;

Best Local Similarity 100.0%; Pred. NO. 27;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DFGWKG 6
|||||
Db 301 dfgwsk 306

Search completed: November 5, 2001, 12:57:31
Job time: 33 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 5, 2001, 12:56:59 ; Search time 12.47 seconds
(without alignments)
10.828 Million cell updates/sec

Title: US-08-894-356C-21

Perfect score: 40

Sequence: 1 DFGWGK 6

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 212252 seqs, 22503292 residues

Total number of hits satisfying chosen parameters: 212252

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 08
Maximum Match 100%

Listing first 45 summaries

Database :

Issued_Patents_AA:*
1: /cgn2_6/prodata/2/1aa/5A.COMB.pep:*
2: /cgn2_6/prodata/2/1aa/5B.COMB.pep:*
3: /cgn2_6/prodata/2/1aa/5A.COMB.pep:*
4: /cgn2_6/prodata/2/1aa/5B.COMB.pep:*
5: /cgn2_6/prodata/2/1aa/PCUTS.COMB.pep:*
6: /cgn2_6/prodata/2/1aa/Backfile1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	40	100.0	7	4	US-09-457-046B-41
2	40	100.0	302	4	US-09-457-046B-8
3	40	100.0	302	4	US-09-457-046B-10
4	40	100.0	302	4	US-09-457-046B-16
5	40	100.0	302	4	US-09-457-046B-18
6	40	100.0	302	4	US-09-457-046B-24
7	40	100.0	303	4	US-09-457-046B-12
8	40	100.0	303	4	US-09-457-046B-20
9	40	100.0	306	4	US-09-457-046B-2
10	40	100.0	306	4	US-09-457-046B-4
11	40	100.0	306	4	US-09-457-046B-22
12	40	100.0	322	4	US-09-457-046B-14
13	40	100.0	331	4	US-09-457-046B-59
14	40	100.0	436	4	US-09-457-046B-62
15	40	100.0	439	4	US-09-457-046B-68
16	40	100.0	448	1	US-08-207-904-2
17	40	100.0	448	1	US-08-207-904-17
18	40	100.0	451	4	US-09-457-046B-69
19	40	100.0	455	4	US-09-457-046B-72
20	40	100.0	460	4	US-09-457-046B-71
21	40	100.0	461	4	US-09-457-046B-64
22	37	92.5	482	4	US-09-457-046B-63
23	35	87.5	301	4	US-09-457-046B-6
24	35	87.5	433	4	US-09-457-046B-66
25	35	87.5	439	4	US-09-457-046B-26
26	35	87.5	440	4	US-09-457-046B-25
27	35	87.5	440	4	US-09-457-046B-45

28	35	87.5	441	4	US-09-457-046B-54	Sequence 54, Appl
29	35	87.5	443	4	US-09-457-046B-50	Sequence 50, Appl
30	35	87.5	445	4	US-09-457-046B-52	Sequence 52, Appl
31	35	87.5	445	4	US-09-457-046B-73	Sequence 73, Appl
32	35	87.5	497	4	US-09-457-046B-65	Sequence 65, Appl
33	31	77.5	45	2	US-08-726-306A-51	Sequence 51, Appl
34	31	77.5	923	3	US-08-936-135-6	Sequence 49, Appl
35	30	75.0	33	2	US-08-031-538-49	Sequence 54, Appl
36	30	75.0	33	2	US-08-031-538-54	Sequence 10, Appl
37	30	75.0	259	1	US-07-981-707-10	Sequence 10, Appl
38	30	75.0	259	1	US-07-988-430-10	Sequence 10, Appl
39	30	75.0	259	1	US-08-425-336-10	Sequence 10, Appl
40	30	75.0	259	1	US-08-488-113B-10	Sequence 10, Appl
41	30	75.0	259	1	US-08-477-484B-10	Sequence 10, Appl
42	30	75.0	259	2	US-08-646-360-10	Sequence 10, Appl
43	30	75.0	259	4	US-08-839-765-10	Sequence 10, Appl
44	30	75.0	259	4	US-09-136-389-10	Sequence 10, Appl
45	30	75.0	259	5	PCT-US92-09487-10	Sequence 10, Appl

ALIGNMENTS

RESULT 1
US-09-457-046B-41
; Sequence 41, Application US/09457046B
; Patent No. 6287835
; GENERAL INFORMATION:
; APPLICANT: Croteau, Rodney et al.
; TITLE OF INVENTION: Transacylases of the Pacifitaxel Biosynthetic Pathway
; FILE REFERENCE: 53679
; CURRENT APPLICATION NUMBER: US/09/457, 046B
; CURRENT FILING DATE: 1999-12-07
; NUMBER OF SEQ ID NOS: 74
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 41
; LENGTH: 7
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:consensus
; OTHER INFORMATION: sequence
US-09-457-046B-41

Query Match 100.0%; Score 40; DB 4; Length 7;
Best Local Similarity 100.0%; Pred. No. 1.6e+05;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 DFGWGK 6
Db 1 DFGWGK 6
RESULT 2
US-09-457-046B-8
; Sequence 8, Application US/09457046B
; Patent No. 6287835
; GENERAL INFORMATION:
; APPLICANT: Croteau, Rodney et al.
; TITLE OF INVENTION: Transacylases of the Pacifitaxel Biosynthetic Pathway
; FILE REFERENCE: 53679
; CURRENT APPLICATION NUMBER: US/09/457, 046B
; CURRENT FILING DATE: 1999-12-07
; NUMBER OF SEQ ID NOS: 74
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 8
; LENGTH: 302
; TYPE: PRT
; ORGANISM: Taxus cuspidata
US-09-457-046B-8

Query Match 100.0%; Score 40; DB 4; Length 302;
Best Local Similarity 100.0%; Pred. No. 9.3;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 DRGWGK 6
|||||
DB 297 DRGWGK 302

RESULT 3
US-09-457-046B-10
; Sequence 10, Application US/09457046B
; Patent No. 6287835

; GENERAL INFORMATION:
; APPLICANT: Croteau, Rodney et al.

; TITLE OF INVENTION: Transcylases of the Facilitaxel Biosynthetic Pathway
; FILE REFERENCE: 53679
; CURRENT APPLICATION NUMBER: US/09/457,046B

; CURRENT FILING DATE: 1999-12-07
; NUMBER OF SEQ ID NOS: 74
; SOFTWARE: Patentln Ver. 2.1

; SEQ ID NO 10
; LENGTH: 302
; TYPE: PRT

; ORGANISM: Taxus cuspidata
US-09-457-046B-10

Query Match 100.0%; Score 40; DB 4; Length 302;
Best Local Similarity 100.0%; Pred. No. 9.3;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 DRGWGK 6
|||||
DB 297 DRGWGK 302

RESULT 4
US-09-457-046B-16
; Sequence 16, Application US/09457046B
; Patent No. 6287835

; GENERAL INFORMATION:
; APPLICANT: Croteau, Rodney et al.

; TITLE OF INVENTION: Transcylases of the Facilitaxel Biosynthetic Pathway
; FILE REFERENCE: 53679
; CURRENT APPLICATION NUMBER: US/09/457,046B

; CURRENT FILING DATE: 1999-12-07
; NUMBER OF SEQ ID NOS: 74
; SOFTWARE: Patentln Ver. 2.1

; SEQ ID NO 16
; LENGTH: 302
; TYPE: PRT

; ORGANISM: Taxus cuspidata
US-09-457-046B-16

Query Match 100.0%; Score 40; DB 4; Length 302;
Best Local Similarity 100.0%; Pred. No. 9.3;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 DRGWGK 6
|||||
DB 297 DRGWGK 302

RESULT 5
US-09-457-046B-18
; Sequence 18, Application US/09457046B
; Patent No. 6287835

; GENERAL INFORMATION:
; APPLICANT: Croteau, Rodney et al.

; TITLE OF INVENTION: Transcylases of the Facilitaxel Biosynthetic Pathway
; FILE REFERENCE: 53679

; CURRENT APPLICATION NUMBER: US/09/457,046B
; CURRENT FILING DATE: 1999-12-07
; NUMBER OF SEQ ID NOS: 74

; SOFTWARE: Patentln Ver. 2.1
; SEQ ID NO 18

; LENGTH: 302
; TYPE: PRT

; ORGANISM: Taxus cuspidata
; FEATURE:

; NAME/KEY: VARIANT
; LOCATION: 164

; OTHER INFORMATION: Any amino acid
US-09-457-046B-18

Query Match 100.0%; Score 40; DB 4; Length 302;
Best Local Similarity 100.0%; Pred. No. 9.3;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 DRGWGK 6
|||||
DB 297 DRGWGK 302

RESULT 6
US-09-457-046B-24
; Sequence 24, Application US/09457046B
; Patent No. 6287835

; GENERAL INFORMATION:
; APPLICANT: Croteau, Rodney et al.

; TITLE OF INVENTION: Transcylases of the Facilitaxel Biosynthetic Pathway
; FILE REFERENCE: 53679
; CURRENT APPLICATION NUMBER: US/09/457,046B

; CURRENT FILING DATE: 1999-12-07
; NUMBER OF SEQ ID NOS: 74
; SOFTWARE: Patentln Ver. 2.1

; SEQ ID NO 24
; LENGTH: 302
; TYPE: PRT

; ORGANISM: Taxus cuspidata
US-09-457-046B-24

Query Match 100.0%; Score 40; DB 4; Length 302;
Best Local Similarity 100.0%; Pred. No. 9.3;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 DRGWGK 6
|||||
DB 297 DRGWGK 302

RESULT 7
US-09-457-046B-12
; Sequence 12, Application US/09457046B
; Patent No. 6287835

; GENERAL INFORMATION:
; APPLICANT: Croteau, Rodney et al.

; TITLE OF INVENTION: Transcylases of the Facilitaxel Biosynthetic Pathway
; FILE REFERENCE: 53679
; CURRENT APPLICATION NUMBER: US/09/457,046B

; CURRENT FILING DATE: 1999-12-07
; NUMBER OF SEQ ID NOS: 74
; SOFTWARE: Patentln Ver. 2.1

; SEQ ID NO 12
; LENGTH: 303
; TYPE: PRT

; ORGANISM: Taxus cuspidata
US-09-457-046B-12

Query Match 100.0%; Score 40; DB 4; Length 303;
Best Local Similarity 100.0%; Pred. No. 9.3;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 DFGWCK 6
111111
DB 298 DFGWCK 303

RESULT 8

US-09-457-046B-20
; Sequence 20, Application US/09457046B
; Patent No. 6287835
; GENERAL INFORMATION:
; APPLICANT: Croteau, Rodney et al.
; TITLE OF INVENTION: Transacylases of the Pacilitaxel Biosynthetic Pathway
; FILE REFERENCE: 53679
; CURRENT APPLICATION NUMBER: US/09/457,046B
; CURRENT FILING DATE: 1999-12-07
; NUMBER OF SEQ ID NOS: 74
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 20
; LENGTH: 303
; TYPE: PRT
; ORGANISM: Taxus cuspidata
US-09-457-046B-20

Query Match 100.0%; Score 40; DB 4; Length 303;

Best Local Similarity 100.0%; Pred. No. 9.3; Mismatches 0; Indels 0; Gaps 0;

OY 1 DFGWCK 6
111111
DB 298 DFGWCK 303

RESULT 9

US-09-457-046B-2
; Sequence 2, Application US/09457046B
; Patent No. 6287835
; GENERAL INFORMATION:
; APPLICANT: Croteau, Rodney et al.
; TITLE OF INVENTION: Transacylases of the Pacilitaxel Biosynthetic Pathway
; FILE REFERENCE: 53679
; CURRENT APPLICATION NUMBER: US/09/457,046B
; CURRENT FILING DATE: 1999-12-07
; NUMBER OF SEQ ID NOS: 74
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 306
; TYPE: PRT
; ORGANISM: Taxus cuspidata
US-09-457-046B-2

Query Match 100.0%; Score 40; DB 4; Length 306;

Best Local Similarity 100.0%; Pred. No. 9.4; Mismatches 0; Indels 0; Gaps 0;

OY 1 DFGWCK 6
111111
DB 301 DFGWCK 306

RESULT 10

US-09-457-046B-4
; Sequence 4, Application US/09457046B
; Patent No. 6287835
; GENERAL INFORMATION:
; APPLICANT: Croteau, Rodney et al.
; TITLE OF INVENTION: Transacylases of the Pacilitaxel Biosynthetic Pathway
; FILE REFERENCE: 53679
; CURRENT APPLICATION NUMBER: US/09/457,046B
; CURRENT FILING DATE: 1999-12-07

; NUMBER OF SEQ ID NOS: 74
; SOFTWARE: PatentIn Ver. 2.1
+; SEQ ID NO 4
; LENGTH: 306
; TYPE: PRT
; ORGANISM: Taxus cuspidata
US-09-457-046B-4

Query Match 100.0%; Score 40; DB 4; Length 306;

Best Local Similarity 100.0%; Pred. No. 9.4; Mismatches 0; Indels 0; Gaps 0;

OY 1 DFGWCK 6
111111
DB 301 DFGWCK 306

RESULT 11

US-09-457-046B-22
; Sequence 22, Application US/09457046B
; Patent No. 6287835
; GENERAL INFORMATION:
; APPLICANT: Croteau, Rodney et al.
; TITLE OF INVENTION: Transacylases of the Pacilitaxel Biosynthetic Pathway
; FILE REFERENCE: 53679
; CURRENT APPLICATION NUMBER: US/09/457,046B
; CURRENT FILING DATE: 1999-12-07
; NUMBER OF SEQ ID NOS: 74
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 22
; LENGTH: 306
; TYPE: PRT
; ORGANISM: Taxus cuspidata
US-09-457-046B-22

Query Match 100.0%; Score 40; DB 4; Length 306;

Best Local Similarity 100.0%; Pred. No. 9.4; Mismatches 0; Indels 0; Gaps 0;

OY 1 DFGWCK 6
111111
DB 301 DFGWCK 306

RESULT 12

US-09-457-046B-14
; Sequence 14, Application US/09457046B
; Patent No. 6287835
; GENERAL INFORMATION:
; APPLICANT: Croteau, Rodney et al.
; TITLE OF INVENTION: Transacylases of the Pacilitaxel Biosynthetic Pathway
; FILE REFERENCE: 53679
; CURRENT APPLICATION NUMBER: US/09/457,046B
; CURRENT FILING DATE: 1999-12-07
; NUMBER OF SEQ ID NOS: 74
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 14
; LENGTH: 322
; TYPE: PRT
; ORGANISM: Taxus cuspidata
US-09-457-046B-14

Query Match 100.0%; Score 40; DB 4; Length 322;

Best Local Similarity 100.0%; Pred. No. 9.8; Mismatches 0; Indels 0; Gaps 0;

OY 1 DFGWCK 6
111111
DB 317 DFGWCK 322

RESULT 13
US-09-457-046B-59
; Sequence 59, Application US/09457046B
; Patent No. 6287835
; GENERAL INFORMATION:
; APPLICANT: Croteau, Rodney et al.
; TITLE OF INVENTION: Transacylases of the Facilitaxel Biosynthetic Pathway
; FILE REFERENCE: 53679
; CURRENT APPLICATION NUMBER: US/09/457,046B
; CURRENT FILING DATE: 1999-12-07
; NUMBER OF SEQ ID NOS: 74
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 59
; LENGTH: 331
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
US-09-457-046B-59

Query Match 100.0%; Score 40; DB 4; Length 331;
Best Local Similarity 100.0%; Pred. No. 10;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 DFGWCK 6
|||||
Db 265 DFGWCK 270

RESULT 14
US-09-457-046B-62
; Sequence 62, Application US/09457046B
; Patent No. 6287835
; GENERAL INFORMATION:
; APPLICANT: Croteau, Rodney et al.
; TITLE OF INVENTION: Transacylases of the Facilitaxel Biosynthetic Pathway
; FILE REFERENCE: 53679
; CURRENT APPLICATION NUMBER: US/09/457,046B
; CURRENT FILING DATE: 1999-12-07
; NUMBER OF SEQ ID NOS: 74
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 62
; LENGTH: 436
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
US-09-457-046B-62

Query Match 100.0%; Score 40; DB 4; Length 436;
Best Local Similarity 100.0%; Pred. No. 13;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 DFGWCK 6
|||||
Db 372 DFGWCK 377

RESULT 15
US-09-457-046B-68
; Sequence 68, Application US/09457046B
; Patent No. 6287835
; GENERAL INFORMATION:
; APPLICANT: Croteau, Rodney et al.
; TITLE OF INVENTION: Transacylases of the Facilitaxel Biosynthetic Pathway
; FILE REFERENCE: 53679
; CURRENT APPLICATION NUMBER: US/09/457,046B
; CURRENT FILING DATE: 1999-12-07
; NUMBER OF SEQ ID NOS: 74
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 68
; LENGTH: 439
; TYPE: PRT
; ORGANISM: Catharanthus roseus

US-09-457-046B-68

Query Match 100.0%; Score 40; DB 4; Length 439;
Best Local Similarity 100.0%; Pred. No. 13;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 DFGWCK 6
|||||
Db 380 DFGWCK 385

Search completed: November 5, 2001, 12:58:09
Job time: 70 sec

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 5, 2001, 12:56:59 ; Search time 12.79 Seconds
(without alignments)
35.735 Million cell updates/sec

Title: US-08-894-356c-21

Perfect score: 40

Sequence: 1 DFGWGK 6

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 219241 seqs, 76174552 residues

Total number of hits satisfying chosen parameters: 219241

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_F68:*
1: PIR1:*
2: PIR2:*
3: PIR3:*
4: PIR4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	40	100.0	436	2 D86378	protein F21J9.8 (1
2	40	100.0	443	2 T45574	anthranilate N-hyd
3	40	100.0	450	2 T45573	anthranilate N-hyd
4	40	100.0	451	2 T00527	hypothetical prote
5	40	100.0	455	2 T09666	probable anthranil
6	40	100.0	460	2 T03274	hsr201 protein, hy
7	40	100.0	461	2 T00918	hypothetical prote
8	40	100.0	475	2 T45961	anthranilate N-ben
9	37	92.5	399	2 F83484	probable mrs trans
10	37	92.5	482	2 G84823	probable anthocyan
11	36	90.0	430	2 T46216	hypothetical prote
12	35	87.5	439	2 T52321	taxadienol acetyl
13	35	87.5	440	2 T52320	10-deacetylacetyl
14	35	87.5	442	2 T10718	anthranilate N-ben
15	35	87.5	445	2 T10717	anthranilate N-ben
16	35	87.5	445	2 T10719	anthranilate N-ben
17	35	87.5	446	2 T10711	anthranilate N-ben
18	35	87.5	464	2 G86453	YUP8H12R.39. homol
19	35	87.5	476	2 D84505	probable membrane
20	35	87.5	480	2 H86411	protein FIK23.12 (
21	35	87.5	485	2 T33865	hypothetical prote
22	35	87.5	572	2 T01056	hypothetical prote
23	34	85.0	113	2 S30515	wound-induced prot
24	34	85.0	129	2 S60978	probable membrane
25	34	85.0	216	2 C85327	hypothetical prote
26	34	85.0	216	2 T02913	hypothetical prote
27	34	85.0	303	2 B83336	hypothetical prote
28	34	85.0	493	2 G84594	probable diacylgly
29	34	85.0	519	2 S54300	transketolase (EC

30	34	85.0	667	2 G69723	transketolase (EC
31	34	85.0	741	2 G84888	probable transkeo
32	34	85.0	754	2 T47886	transketolase-like
33	33	82.5	730	2 A75486	hypothetical prote
34	32	80.0	234	2 T36369	response regulator
35	32	80.0	281	2 A86742	hypothetical prote
36	32	80.0	299	2 I36787	hypothetical prote
37	32	80.0	346	2 T10173	24-sterol C-methyl
38	32	80.0	367	2 T06780	probable 24-sterol
39	32	80.0	409	2 T19688	hypothetical prote
40	32	80.0	428	2 T48008	hypothetical prote
41	31	77.5	106	2 B64350	hypothetical prote
42	31	77.5	110	2 F71129	hypothetical prote
43	31	77.5	110	2 F75034	hypothetical prote
44	31	77.5	112	2 J00151	biphenyl-2,3-diol
45	31	77.5	127	2 D82621	hypothetical prote

ALIGNMENTS

RESULT 1
D86378
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 31-Mar-2001
C:Accession: D86378
R:Phenology: A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, ansen, N.F.; Hughes, B.; Hultzer, L.
Nature 408, 816-820, 2000
A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Malt, R.; Marzia A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallo ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
A:Reference number: A86141; MUID:21016719
A:Accession: D86378
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-436 <STO>
A:Cross-references: GB:AE005172; NID:g9743331; PIDN:AAF97955.1; GSPDB:GN00141
C:Genetics:
A:Gene: F21J9.8
A:Map position: 1

Query Match 100.0%; Score 40; DB 2; Length 436;
Best Local Similarity 100.0%; Pred. No. 11;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DFGWGK 6
DB 372 DFGWGK 377

RESULT 2
T45574
N:Alternate names: protein F1IC1.120
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 04-Feb-2000 #sequence_revision 04-Feb-2000 #text_change 04-Feb-2000
C:Accession: T45574
K:Barques, M.; Collado, M.C.; Navarro, P.; Terol, J.; Perez-Alonso, M.; Mewes, H.W.; submitted to the Protein Sequence Database, December 1999
A:Reference number: Z23007
A:Accession: T45574
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-443 <BAR>
A:Cross-references: EMBL:AL132976
A:Experimental source: Cultivar Columbia; BAC clone F1IC1

C:Genetics:
A:Map position: 3
A>Note: FLIC1.120

Query Match 100.0%: Score 40; DB 2; Length 443;
Best Local Similarity 100.0%: Pred. No. 11;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 DFGWCK 6
|||||
DB 386 DFGWCK 391

RESULT 3
T45573
anthranilate N-hydroxycinnamoyl/benzoyltransferase-like protein - Arabidopsis thaliana
N:Alternate names: protein FLIC1.110
C:Species: Arabidopsis thaliana (mouse-ear cress)
C>Date: 04-Feb-2000 #sequence_revision 04-Feb-2000 #text_change 04-Feb-2000
C:Accession: T45573
R:Barques, M.; Collado, M.C.; Navarro, P.; Terol, J.; Perez-Alonso, M.; Mewes, H.W.; May
submitted to the Protein Sequence Database, December 1999
A:Reference number: 223007
A:Accession: T45573
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-450 <BAR>
A:Cross-references: EMBL:AL132976
A:Experimental source: cultivar Columbia; BAC clone FLIC1
C:Genetics:
A:Map position: 3
A>Note: FLIC1.110

Query Match 100.0%: Score 40; DB 2; Length 450;
Best Local Similarity 100.0%: Pred. No. 11;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 DFGWCK 6
|||||
DB 391 DFGWCK 396

RESULT 4
T00527
hypothetical protein At2g19070 [imported] - Arabidopsis thaliana
N:Alternate names: hypothetical protein T20K24.8
C:Species: Arabidopsis thaliana (mouse-ear cress)
C>Date: 01-Feb-1999 #sequence_revision 01-Feb-1999 #text_change 02-Feb-2001
C:Accession: T00527; B84572
R:Rounsley, S.D.; Kaul, X.; Ketchum, K.A.; Crosby, M.L.; Brandon, R.C.; Sykes,
submitted to the EMBL Data Library, July 1997
A:Description: Arabidopsis thaliana chromosome II BAC T20K24 genomic sequence.
A:Reference number: Z14167
A:Accession: T00527
A:Status: translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-451 <ROU>
A:Cross-references: EMBL:AC002392; NID:g3176701; PID:g3176709
A:Experimental source: cultivar Columbia
R:Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.;
M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanKen, S.E.; Umayam, L.; Tallon, L.;
Euss, D.; Nierman, W.C.; White, O.; Elsen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J.
Nature 402, 761-768, 1999
A>Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.
A:Reference number: A84420; MUID:20083487
A:Accession: B84572
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-451 <STO>
A:Cross-references: GB:AE002093; NID:g3176709; PIDN:AAD12025.1; GSPDB:GND0139
C:Genetics:

A:Gene: At2g19070; T20K24.8
A:Map position: 2
A:Introns: 322/1

Query Match 100.0%: Score 40; DB 2; Length 451;
Best Local Similarity 100.0%: Pred. No. 11;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 DFGWCK 6
|||||
DB 397 DFGWCK 402

RESULT 5
T09666
probable anthranilate N-benzoyltransferase (EC 2.3.1.144) - muskmelon (fragment)
N:Alternate names: anthranilate N-hydroxycinnamoyl/benzoyltransferase
C:Species: Cucumis melo (muskmelon)
C>Date: 16-Jul-1999 #sequence_revision 16-Jul-1999 #text_change 21-Jul-2000
C:Accession: T09666
R:Aggelis, A.; John, I.; Karvouni, Z.; Grierson, D.
Plant Mol. Biol. 33, 313-322, 1997
A>Title: Characterization of two cDNA clones for mRNAs expressed during ripening of m
A:Reference number: Z16810; MUID:971188564
A:Accession: T09666
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-455 <AGG>
A:Cross-references: EMBL:Z70521; NID:g1843439; PIDN:CAA94432.1; PID:g1843440
A:Experimental source: cultivar Cantaloupe charentais; pericarp of ripe fruit
C:Keywords: acyltransferase; coenzyme A

Query Match 100.0%: Score 40; DB 2; Length 455;
Best Local Similarity 100.0%: Pred. No. 11;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 DFGWCK 6
|||||
DB 374 DFGWCK 379

RESULT 6
T03274
hsr201 protein, hypersensitivity-related - common tobacco
C:Species: Nicotiana tabacum (common tobacco)
C>Date: 24-Mar-1999 #sequence_revision 24-Mar-1999 #text_change 21-Jul-2000
C:Accession: T03274
R:Czernik, P.; Huang, H.C.; Marco, Y.
Plant Mol. Biol. 31, 255-265, 1996
A>Title: Characterization of hsr201 and hsr515, two tobacco genes preferentially expr
A:Reference number: Z14876; MUID:96343929
A:Accession: T03274
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-460 <CZE>
A:Cross-references: EMBL:X95343; NID:g1171576; PIDN:CAA64636.1; PID:g1171577
A:Experimental source: cultivar bottom special
C:Genetics:
A:Gene: hsr201

Query Match 100.0%: Score 40; DB 2; Length 460;
Best Local Similarity 100.0%: Pred. No. 11;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 DFGWCK 6
|||||
DB 382 DFGWCK 387

RESULT 7

T00918
hypothetical protein F21B7.32 - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 12-Feb-1999 #sequence_revision 12-Feb-1999 #text_change 22-Oct-1999
C:Accession: T00918
R:Shinn, P.; Buehler, E.; Dewar, K.; Feng, J.; Kim, C.; Li, Y.; Sun, H.; Conway, A.; Cor
eologis, A.; Ecker, J.R.
Submitted to the EMBL Data Library, January 1998
A:Description: Genomic sequence for Arabidopsis thaliana BAC F21B7.
A:Reference number: Z14208
A:Accession: T00918
A>Status: translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-461 <SH1>
A:Cross-references: EMBL:AC002560; NID:g2618677; PID:g2809263; GSPDB:GN00059; ATSP:F21B7
C:Genetics:
A:Gene: ATSP:F21B7.32
A:Map position: 1

Query Match 100.0%; Score 40; DB 2; Length 461;
Best Local Similarity 100.0%; Pred. No. 11;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 DFGWCK 6
|||||
Db 400 DFGWCK 405

RESULT 8
T45961
anthranilate N-benzoyltransferase-like protein - Arabidopsis thaliana
N:Alternate names: protein F7J8.190
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 04-Feb-2000 #sequence_revision 04-Feb-2000 #text_change 04-Feb-2000
C:Accession: T45961
R:Bevan, M.; Zimmermann, W.; Gruenisen, A.; Mambutt, R.; Bancroft, I.; Mewes, H.W.; Lem
Submitted to the Protein Sequence Database, January 2000
A:Reference number: Z23018
A:Accession: T45961
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-475 <BEV>
A:Cross-references: EMBL:ALJ37189
A:Experimental source: cultivar Columbia; BAC clone F7J8
C:Genetics:
A:Map position: 5
A:Note: F7J8.190

Query Match 100.0%; Score 40; DB 2; Length 475;
Best Local Similarity 100.0%; Pred. No. 11;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 DFGWCK 6
|||||
Db 411 DFGWCK 416

RESULT 9
F83484
Probable MFS transporter PA1286 [imported] - Pseudomonas aeruginosa (strain PA01)
C:Species: Pseudomonas aeruginosa
C:Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 31-Dec-2000
C:Accession: F83484
R:Stover, C.K.; Pham, X.O.; Erwin, A.L.; Mizoguchi, S.D.; Warren, P.; Hickey, M.J.; B
adman, S.; Vran, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Lapid, K.; Lim,
.; Lory, S.; Olson, M.V.
Nature 406, 959-964, 2000
A:Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic patho
A:Reference number: AB2950; MUID:20437337
A:Accession: F83484
A>Status: preliminary

A:Molecule type: DNA
A:Residues: 1-399 <STO>
A:Cross-references: GB:AE004558; GB:AE004091; NID:g9947217; PIDN:AA04675.1; GSPDB:GN
A:Experimental source: strain PA01
C:Genetics:
A:Gene: PA1286

Query Match 92.5%; Score 37; DB 2; Length 399;
Best Local Similarity 83.3%; Pred. No. 32;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 DFGWCK 6
|||||
Db 36 DFGWCK 41

RESULT 10
G84823
probable anthocyanin 5-aromatic acyltransferase [imported] - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 02-Feb-2001
C:Accession: G84823
R:Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y
M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.E.; Umayam, L.; Tallon
euss, D.; Niernman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter
Nature 402, 761-768, 1999
A:Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.
A:Reference number: AB4420; MUID:20083487
A:Accession: G84823
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-482 <STO>
A:Cross-references: GB:AE002093; NID:g2086651; PIDN:AAB95283.1; GSPDB:GN00139
C:Genetics:
A:Gene: AT2G39980
A:Map position: 2

Query Match 92.5%; Score 37; DB 2; Length 482;
Best Local Similarity 83.3%; Pred. No. 38;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 DFGWCK 6
|||||
Db 422 DFGWCK 427

RESULT 11
T46216
hypothetical protein T8P19.230 - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 04-Feb-2000 #sequence_revision 04-Feb-2000 #text_change 17-Mar-2000
C:Accession: T46216
R:Choisne, N.; Robert, C.; Brottier, P.; Wincker, P.; Cattivello, L.; Artiguenave, F.;
Submitted to the Protein Sequence Database, December 1999
A:Reference number: Z23008
A:Accession: T46216
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-430 <CHO>
A:Cross-references: EMBL:ALJ33315
A:Experimental source: cultivar Columbia; BAC clone T8P19
C:Genetics:
A:Map position: 3
A:Introns: 144/3
A:Note: T8P19.230
C:Superfamily: Arabidopsis CER2 protein

Query Match 90.0%; Score 36; DB 2; Length 430;
Best Local Similarity 83.3%; Pred. No. 51;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 DFGWG 6
 |||||
 Db 378 DFGWG 383

RESULT 12

T52321
 taxadienol acetyl transferase [imported] - Taxus cuspidata
 C:Species: Taxus cuspidata
 C:Date: 20-Oct-2000 #sequence_revision 20-Oct-2000 #text_change 20-Oct-2000
 C:Accession: T52321
 R:Walker, K.; Schendorf, A.; Croteau, R.
 Submitted to the EMBL Data Library, September 1999
 A:Description: Molecular Cloning of a Taxa-4(20),11(12)-dien-5(alpha)-ol-O-Acetyl Transf
 A:Reference number: 226036
 A:Accession: T52321
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: mRNA
 A:Residues: 1-439 <MAL>
 A:Cross-references: EMBL:AF190130; PIDN:AAF34254.1
 C:Genetics:
 A:Gene: TAT

Query Match 87.5%; Score 35; DB 2; Length 439;
 Best Local Similarity 100.0%; Pred. No. 77;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DFGWG 5
 |||||
 Db 373 DFGWG 377

RESULT 13

T52320
 10-deacetyl baccatin III-10-O-acetyl transferase [imported] - Taxus cuspidata
 C:Species: Taxus cuspidata
 C:Date: 20-Oct-2000 #sequence_revision 20-Oct-2000 #text_change 20-Oct-2000
 C:Accession: T52320
 R:Walker, K.; Croteau, R.
 Proc. Natl. Acad. Sci. U.S.A. 97, 583-587, 2000
 A:Title: Molecular cloning of a 10-deacetyl baccatin III-10-O-acetyl transferase cDNA fr
 A:Reference number: 226035
 A:Accession: T52320
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: mRNA
 A:Residues: 1-440 <MAL>
 A:Cross-references: EMBL:AF193765; PIDN:AAF27621.1
 C:Genetics:
 A:Gene: DBAT

Query Match 87.5%; Score 35; DB 2; Length 440;
 Best Local Similarity 100.0%; Pred. No. 77;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DFGWG 5
 |||||
 Db 372 DFGWG 376

RESULT 14

T10718
 anthranilate N-benzoyltransferase (EC 2.3.1.144) (clone pchcbt1) - clove pink (fragment
 N:Alternate names: anthranilate N-hydroxycinnamoyl/benzoyltransferase
 C:Species: Dianthus caryophyllus (clove pink)
 C:Date: 16-Jul-1999 #sequence_revision 16-Jul-1999 #text_change 21-Jul-2000
 C:Accession: T10718
 R:Yang, Q.; Reinhard, K.; Schiltz, E.; Matern, U.
 Plant Mol. Biol. 35, 777-789, 1997
 A:Title: Characterization and heterologous expression of hydroxycinnamoyl/ benzoyl-CoA:
 lus L.

A:Reference number: 217095; MUID:98088004
 A:Accession: T10718
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: mRNA
 A:Residues: 1-442 <YAN>

A:Cross-references: EMBL:Z84384; NID:g2239084; PIDN:CAB06428.1; PID:g2239085
 C:Function:
 A:Description: catalyzes the synthesis of anthranilate
 A:Pathway: phytoalexin biosynthesis
 C:Keywords: acyltransferase; coenzyme A

Query Match 87.5%; Score 35; DB 2; Length 442;
 Best Local Similarity 100.0%; Pred. No. 78;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DFGWG 5
 |||||
 Db 389 DFGWG 393

RESULT 15

T10717
 anthranilate N-benzoyltransferase (EC 2.3.1.144) (clone pchcbt1) - clove pink
 N:Alternate names: anthranilate N-hydroxycinnamoyl/benzoyltransferase
 C:Species: Dianthus caryophyllus (clove pink)
 C:Date: 16-Jul-1999 #sequence_revision 16-Jul-1999 #text_change 21-Jul-2000
 C:Accession: T10717
 R:Yang, Q.; Reinhard, K.; Schiltz, E.; Matern, U.
 Plant Mol. Biol. 35, 777-789, 1997
 A:Title: Characterization and heterologous expression of hydroxycinnamoyl/ benzoyl-Co
 lus L.

A:Reference number: 217095; MUID:98088004
 A:Accession: T10717
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: mRNA
 A:Residues: 1-445 <YAN>
 A:Cross-references: EMBL:Z84383; NID:g2239082; PIDN:CAB06427.1; PID:g2239083
 C:Function:
 A:Description: catalyzes the synthesis of anthranilate
 A:Pathway: phytoalexin biosynthesis
 C:Keywords: acyltransferase; coenzyme A

Query Match 87.5%; Score 35; DB 2; Length 445;
 Best Local Similarity 100.0%; Pred. No. 78;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DFGWG 5
 |||||
 Db 392 DFGWG 396

Search completed: November 5, 2001, 12:57:51
 Job time: 52 sec

THIS PAGE BLANK (uspto)

DR EMBL: M65140; AAA51661.1; -
 DR HSSP: P19367; HKB.
 DR InterPro: IPR001312; Hexokinase.
 DR Pfam: PF00349; hexokinase; 2.
 DR PRINTS: PR00475; HEXOKINASE.
 DR PRODOM: PD001109; Hexokinase; 2.
 DR PROSITE: PS00378; HEXOKINASES; 2.
 DR Transferrase; Kinase; Glycolysis; Allosteric enzyme; Repeat;
 KW ATP-binding; Membrane.
 FT DOMAIN 1 12 HYDROPHOBIC.
 FT DOMAIN 13 475 REGULATORY (BY SIMILARITY).
 FT DOMAIN 476 918 CATALYTIC (BY SIMILARITY).
 FT DOMAIN 149 175 GLUCOSE-BINDING (POTENTIAL).
 FT DOMAIN 597 623 GLUCOSE-BINDING (POTENTIAL).
 FT NP_BIND 84 89 ATP (POTENTIAL).
 FT NP_BIND 532 537 ATP (POTENTIAL).
 FT BINDING 558 558 ATP (POTENTIAL).
 SQ SEQUENCE 918 AA; 103064 MW; IDCFF7FID06FE2B6 CRC64;

Query Match 87.5%; Score 35; DB 1; Length 918;
 Best Local Similarity 100.0%; Pred. No. 69;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 DFGWG 5
 |||||
 Db 246 DFGWG 250

RESULT 2
 Y052_HUMAN STANDARD; PRT; 1045 AA.
 AC P42285;
 DT 01-NOV-1995 (Rel. 32, Created)
 DT 01-NOV-1995 (Rel. 32, Last sequence update)
 DT 20-AUG-2001 (Rel. 40, Last annotation update)
 DE HYPOTHETICAL PROTEIN KIAA0052 (FRAGMENT).
 GN KIAA0052.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=bone marrow;
 RX MEDLINE=96051398; PubMed=7584044;
 RA Nomura N., Nagase T., Miyajima N., Sazuka T., Sato S.,
 RA Seki N., Kawarabayashi Y., Ishikawa K.-I., Tabata S.;
 RT "Prediction of the coding sequences of unidentified human genes. II.
 RT The coding sequences of 40 new genes (KIAA0041-KIAA0080) deduced by
 RT analysis of cDNA clones from human cell line KG-1.";
 RL DINA Res. 1:223-229(1994).
 CC -1- SUBCELLULAR LOCATION: NUCLEAR (POTENTIAL).
 CC -1- SIMILARITY: BELONGS TO THE SKI2 SUBFAMILY OF HELICASES.
 CC -1- CAUTION: THIS IS A CONCEPTUAL TRANSLATION: A FRAMESHIFT WAS
 CC INTRODUCED IN POSITION 837 TO MAKE THE PROTEIN LONGER AND MAXIMIZE
 CC THE SIMILARITY WITH SKI2.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC
 CC EMBL: D29641; BAA06124.1; ALT_FRAME.
 CC InterPro: IPR001410; DEAD.
 CC InterPro: IPR001650; Helicase_C.
 CC Pfam: PF00270; DEAD; 1.
 CC Pfam: PF00271; helicase_C; 1.
 CC SMART: SM00487; DEXDC; 1.
 CC SMART: SM00490; HELIC; 1.

KW Hypothetical protein; Helicase; ATP-binding; Nuclear protein.
 FT NON_TER 1 1
 FT NP_BIND 165 172 ATP (POTENTIAL).
 FT SITE 256 259 DEH BOX.
 FT SEQUENCE 1045 AA; 118243 MW; 7B16FF8E78049C20 CRC64;

Query Match 87.5%; Score 35; DB 1; Length 1045;
 Best Local Similarity 100.0%; Pred. No. 77;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 DFGWG 5
 |||||
 Db 672 DFGWG 676

RESULT 3
 YN09_YEAST STANDARD; PRT; 129 AA.
 AC P53903;
 DT 01-OCT-1996 (Rel. 34, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 15-DEC-1998 (Rel. 37, Last annotation update)
 DE HYPOTHETICAL 15.1 KDA PROTEIN IN RPC8-MEA2 INTERGENIC REGION.
 GN YNL149C OR N1774.
 OS Saccharomyces cerevisiae (Baker's yeast).
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 OC Saccharomycetales; Saccharomycetaceae; Saccharomyces.
 OX NCBI_TaxID=4932;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=5288C / FY1679;
 RX MEDLINE=96287653; PubMed=8666380;
 RA Nasr F., Becam A.-M., Herbert C.J.;
 RT "The sequence of 36.8 kb from the left arm of chromosome XIV reveals
 RT 24 complete open reading frames: 18 correspond to new genes, one of
 RT which encodes a protein similar to the human myotonic dystrophy
 RT kinase.";
 RL Yeast 12:169-175(1996).
 RN [2]
 RP SEQUENCE OF 1-65 FROM N.A.
 RC STRAIN=5288C;
 RX MEDLINE=96109932; PubMed=8619318;
 RA Maillet L., Bussiereau F., Jacquet M.;
 RT "A 43.5 kb segment of yeast chromosome XIV, which contains MEA2,
 RT MEP2, CAP/SRV2, NAM9, FRB1/PPRI/RBP1, MOW22 and Cpy1, predicts an
 RT adenosine deaminase gene and 14 new open reading frames.";
 RL Yeast 11:1195-1209(1995).

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC
 CC EMBL: X92517; CAA63290.1; -
 CC DR EMBL: Z71426; CAA96037.1; -
 CC DR EMBL: Z71424; CAA96033.1; -
 CC DR SGD: S0005093; YNL149C.
 KW Hypothetical protein.
 SQ SEQUENCE 129 AA; 15053 MW; A1A4E3DA8CEAA06 CRC64;

Query Match 85.0%; Score 34; DB 1; Length 129;
 Best Local Similarity 100.0%; Pred. No. 19;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 FCGWK 6
 |||||
 Db 84 FCGWK 88


```

AC P45694:
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DE 20-AUG-2001 (Rel. 40, Last annotation update)
DE TRANSEKETOYLASE (EC 2.2.1.1).
GN TKT OR TKTa.
OS Bacillus subtilis.
OC Bacteria, Firmicutes, Bacillus/Clostridium group;
OC Bacillus/Staphylococcus group; Bacillus.
OX NCBI_TaxID=1423;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=168;
RA Rose M., Eutian K.;
RL Submitted (JUN-1996) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE OF 561-667 FROM N.A.
RC STRAIN=168;
RA Schott T., von Machenfeldt C., Hederstedt L.;
RL Submitted (JUN-1995) to the EMBL/GenBank/DBJ databases.
CC -1- CATALYTIC ACTIVITY: SEDOHEPTULOSE 7-PHOSPHATE + D-GLYCERALDEHYDE
CC 3-PHOSPHATE -> D-RIBOSE 5-PHOSPHATE + D-XULULOSE 5-PHOSPHATE.
CC -1- COPACTOR: THIAMINE PYROPHOSPHATE.
CC -1- SIMILARITY: BELONGS TO THE TRANSEKETOYLASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: Z73234; CAA97616.1; -
DR EMBL: X87845; CAA61113.1; -
DR EMBL: Z59113; CAA13673.1; -
DR HSSP: P23254; IAYO;
DR Subtilist; BG11247; Tkt.
DR InterPro: IPR000360; Transketolase.
DR Pfam: PF00456; transketolase; 1.
DR PROSITE: PS00801; TRANSEKETOYLASE_1; 1.
DR PROSITE: PS00802; TRANSEKETOYLASE_2; 1.
DR Transferrase: Thiamine pyrophosphate; Complete proteome.
KW SEQUENCE 667 AA; 72344 MW; D93BCACD246148AF CRC64;
SQ
Query Match 85.0%; Score 34; DB 1; Length 667;
Best Local Similarity 100.0%; Pred. No. 78;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 2 FCGWCK 6
DB 619 FCGWCK 623

```

```

RT "Channel catfish virus: a new type of herpesvirus.";
RL Virology 186:9-14(1992).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: M75136; AAA88120.1; -
DR Figs; 136787; 136787.
DR Hypothetical protein.
KW SEQUENCE 298 AA; 32399 MW; 4C90B1DEEA203FA9 CRC64;
SQ
Query Match 80.0%; Score 32; DB 1; Length 298;
Best Local Similarity 66.7%; Pred. No. 85;
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
OY 1 FCGWCK 6
DB 67 FCGWCK 72

```

```

RESULT 8
Y402_METUA STANDARD; PRT; 106 AA.
ID Y402_METUA
AC 057845;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DE 20-AUG-2001 (Rel. 40, Last annotation update)
DE HYPOTHETICAL PROTEIN MJ0402.
GN MJ0402.
OS Methanococcus jannaschii.
OC Archaea, Euryarchaeota; Methanococcales; Methanococcaceae;
OC Methanococcus.
OX NCBI_TaxID=2190;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=JAL-1 / DSM 2661 / ATCC 43067;
RX MEDLINE=96337999; PubMed=8688087;
RA Bull C.J., White O., Olsen G.J., Zhou L., Fleischmann R.D.,
RA Sutcliffe G.G., Blake O., Fitzgerald L.M., Clayton R.A., Gocayne J.D.,
RA Kierulff A.R., Dougherty B.A., Tomb J.F., Adams M.D., Reisch C.I.,
RA Overbeek R., Kirkness E.F., Weinstock K.G., Merrick J.M., Glodok A.,
RA Scott J.L., Geoghegan N.S.M., Weidman J.F., Fuhrmann J.L., Nguyen D.,
RA Ullrichbeck T.R., Kelley J.M., Peterson J.D., Sadow P.W., Hanna M.C.,
RA Klenk H.-P., Fraser C.M., Hurst M.A., Kaine B.P., Borodovsky M.,
RA Klenk H.-P., Fraser C.M., Smith H.O., Woese C.R., Venter J.C.;
RT "Complete genome sequence of the methanogenic archaeon, Methanococcus
RT jannaschii."
RL Science 273:1058-1073(1996).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: U67492; AAB98395.1; -
DR TIGR: MJ0402; -
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 106 AA; 12600 MW; 6398888ABD46CF93 CRC64;

```

```

Query Match 77.5%; Score 31; DB 1; Length 106;
Best Local Similarity 80.0%; Pred. No. 52;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

```

OY 2 FGWCK 6
DB 2 FGWGR 6

RESULT 9
Y473_AOUAE STANDARD: PRT; 215 AA.

AC 066771;
DT 20-AUG-2001 (Rel. 40, Created)
DT 20-AUG-2001 (Rel. 40, Last sequence update)
DE HYPOTHETICAL PROTEIN AQ_473.

GN AQ_473.
OS Aquifex aeolicus.
OC Bacteria; Aquificales; Aquificaceae; Aquifex.
OX NCBI_TaxId=63363;

RT
RN
RP SEQUENCE FROM N.A.

RC STRAIN=VF5;
RX MEDLINE=98196666; PubMed=9537320;

RA Deckert G., Warren P.V., Gaasterland T., Young W.G., Lenox A.L.,
Graham D.E., Overbeek R., Sneed M.A., Anujay M., Huber R.,
Feldman R.A., Short J.M., Olson G.J., Swanson R.V.;

RT "The complete genome of the hyperthermophilic bacterium Aquifex
aeolicus".

RT Nature 392:353-358(1998).

CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (POTENTIAL).

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation-
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).

DR EMBL: AF000691; AAC06731.1; -
KW Hypothetical protein; Transmembrane; Complete proteome.

FT TRANSMEM 77 99 POTENTIAL.
FT TRANSMEM 117 139 POTENTIAL.

FT TRANSMEM 152 174 POTENTIAL.
FT TRANSMEM 189 208 POTENTIAL.

SO SEQUENCE 215 AA; 23986 MW; BE87A517459DAECC CRC64;

Query Match 77.5%; Score 31; DB 1; Length 215;
Best Local Similarity 80.0%; Pred. No. 95;

Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 2 FGWCK 6
DB 85 YGWCK 89

RESULT 10
FGFF_MOUSE STANDARD: PRT; 218 AA.

AC 035622;
DT 15-JUL-1999 (Rel. 38, Created)
DT 15-JUL-1999 (Rel. 38, Last sequence update)

DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE FIBROBLAST GROWTH FACTOR-15 PRECURSOR (FGF-15).

GN FGF15.
OS Mus musculus (Mouse).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

OX NCBI_TaxId=10090;

RT
RN
RP SEQUENCE FROM N.A.

RX MEDLINE=97454242; PubMed=9310317;

RA McWhirter J.R., Goulding M., Weiner J.A., Chun J., Murre C.;

RT "A novel fibroblast growth factor gene expressed in the developing

RT nervous system is a downstream target of the chimeric homeodomain
oncoprotein E2A-Pbx1."

CC -1- FUNCTION: MAY PLAY AN IMPORTANT ROLE IN REGULATING CELL DIVISION
AND PATTERNING WITHIN SPECIFIC REGIONS OF THE EMBRYONIC BRAIN.

CC SPINAL CORD AND SENSORY ORGANS.

CC -1- SUBCELLULAR LOCATION: SECRETED.

CC -1- TISSUE SPECIFICITY: EXPRESSED IN THE DEVELOPING BRAIN.

CC -1- SIMILARITY: BELONGS TO THE HEPARIN-BINDING GROWTH FACTORS FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation-
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).

DR EMBL: AF007268; AAB63197.1; -
DR MCD; MG1:1096383; Fgf15.

DR InterPro: IPR002209; HBG_FGF.
DR InterPro: IPR002348; IL1_HBGF.

DR Pfam: PF00167; FGF_1.
DR PRINTS: PR00262; IL1HBGF.

DR PRODOM: PD000831; HBG_FGF; 1.
DR SMART: SM00442; FGF_1.

DR PROSITE: PS00247; HBG_FGF; 1.
KW Growth factor; Signal.

FT SIGNAL 1 25 POTENTIAL.
FT CHAIN 26 218 FIBROBLAST GROWTH FACTOR-15.

SO SEQUENCE 218 AA; 25236 MW; A96B0D771FE125A5 CRC64;

Query Match 77.5%; Score 31; DB 1; Length 218;
Best Local Similarity 80.0%; Pred. No. 95;

Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 2 FGWCK 6
DB 44 YGWCK 48

RESULT 11
OPSR_CANFA STANDARD: PRT; 273 AA.

AC 018914;
DT 15-JUL-1998 (Rel. 36, Created)

DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 20-AUG-2001 (Rel. 40, Last annotation update)

DE RED-SENSITIVE OPSIN (RED CONE PHOTORECEPTOR PIGMENT) (FRAGMENT).

GN OPN1LM OR RCP.

OS Canis familiaris (Dog).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.

OX NCBI_TaxId=9615;

RT
RN
RP SEQUENCE FROM N.A.
RX MEDLINE=98242086; PubMed=9580985;

RA Yokoyama S., Radlwimmer F.B.;

RT "The 'five-sites' rule and the evolution of red and green color
vision in mammals."

CC -1- FUNCTION: VISUAL PIGMENTS ARE THE LIGHT-ABSORBING MOLECULES THAT
MEDIATE VISION. THEY CONSIST OF AN APOPROTEIN, OPSIN, COVALENTLY
LINKED TO CIS-RETINAL.

CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN.

CC -1- TISSUE SPECIFICITY: THE THREE COLOR PIGMENTS ARE FOUND IN THE CONE
PHOTORECEPTOR CELLS.

CC -1- PTM: SOME OR ALL OF THE CARBOXYL-TERMINAL SER OR THR RESIDUES MAY
BE PHOSPHORYLATED.

CC -1- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see <http://www.isb-sdb.ch/announce/> or send an email to license@sdb.ch).

DR EMBL: AF031533; AAB86633.1; -
 DR GCRNB: GCR2500; -
 DR InterPro: IPR000276; GPCR_Rhodopsin.
 DR InterPro: IPR001760; Opsin.
 DR Pfam: PF00001; 7tm_1; 1.
 DR PROSITE: PS00237; G-PROTEIN_RECEPTOR_F1_1; 1.
 DR PROSITE: PS50262; G-PROTEIN_RECEPTOR_F1_2; 1.
 DR PROSITE: PS00238; OPSIN; PARTIAL.
 KW Photoreceptor: Retinal protein; Transmembrane; Glycoprotein; Vision; Phosphorylation; Lipoprotein; Palmitate; G-protein coupled receptor.
 FT DOMAIN 1 1 1
 FT TRANSMM 6 30 EXTRACELLULAR (POTENTIAL).
 FT TRANSMM 31 42 CYTOPLASMIC (POTENTIAL).
 FT TRANSMM 43 68 2 (POTENTIAL).
 FT TRANSMM 69 82 EXTRACELLULAR (POTENTIAL).
 FT TRANSMM 83 102 3 (POTENTIAL).
 FT TRANSMM 103 121 CYTOPLASMIC (POTENTIAL).
 FT TRANSMM 122 145 4 (POTENTIAL).
 FT TRANSMM 146 171 EXTRACELLULAR (POTENTIAL).
 FT TRANSMM 172 199 5 (POTENTIAL).
 FT TRANSMM 200 221 CYTOPLASMIC (POTENTIAL).
 FT TRANSMM 222 245 6 (POTENTIAL).
 FT TRANSMM 246 253 EXTRACELLULAR (POTENTIAL).
 FT TRANSMM 254 273 7 (POTENTIAL).
 FT DISULFD 79 156 POTENTIAL.
 FT BINDING 265 265 RETINAL CHROMOPHORE (BY SIMILARITY).
 FT NON_TER 273 273
 SQ SEQUENCE 273 AA; 30373 MW; C1A27C95D1649F85 CRC64;

Query Match 77.5%; Score 31; DB 1; Length 273;
 Best Local Similarity 80.0%; Pred. No. 1.2e+02;
 Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 2 FEGWK 6
 DB 142 FEGWK 146

RESULT 12
 BHCL_RHGO STANDARD; PRT; 291 AA.

AC P47231;
 DT 01-FEB-1996 (Rel. 33, Created)
 DT 01-FEB-1996 (Rel. 33, Last sequence update)
 DT 15-JUL-1999 (Rel. 38, Last annotation update)
 DE BIPIHENTL-2,3-DIOL 1,2-DIOXYGENASE I (EC 1.13.11.39) (230HBP OXYGENASE I) (2,3-DIHYDROXYBIPIHENTYL DIOXYGENASE I) (DHBD I).
 GN BPHC1.
 OS Rhodococcus globerulus.
 CC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae; Actinomycetales; Corynebacteriaceae; Nocardiaaceae; Rhodococcus.
 CC NCBI_TaxID=33008;
 OX NCBI_TaxID=33008;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=P6;
 RX MEDLINE=94171820; PubMed=8126007;
 RA Asturias J.A., Eltis L.D., Prucha M., Timmis K.N.,
 RT "Analysis of three 2,3-dihydroxybiphenyl 1,2-dioxygenases found in Rhodococcus globerulus P6. Identification of a new family of extradiol dioxygenases.";
 RL J. Biol. Chem. 269:7807-7815(1994).
 CC -I- CATALYTIC ACTIVITY: BIPIHENTL-2,3-DIOL + O(2) = 2-HYDROXY-6-OXO-6-PHENYLHEXA-2,4-DIENATE + H(2)O.

CC -I- COFACTOR: FERROUS ION.
 CC -I- PATHWAY: DEGRADATION OF BIPIHENYLS AND POLYCHLOROBIPHENYLS (PCB) TO BENZOIC ACID AND CHLOROBENZOIC ACIDS.
 CC -I- SIMILARITY: BELONGS TO THE EXTRADIOL RING-CLEAVAGE DIOXYGENASE FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see <http://www.isb-sdb.ch/announce/> or send an email to license@sdb.ch).

DR EMBL: X75633; CAA53297.1; -
 DR HSSP: P47228; 1HAN.
 DR InterPro: IPR000486; Extradiol_dioxygenase.
 DR InterPro: IPR000325; Glyoxalase_1.
 DR Pfam: PF00903; Glyoxalase; 1.
 DR Prodom: PD000977; Extradiol_dioxygenase; 1.
 DR PROSITE: PS00082; EXTRADIOL_DIOXYGENAS; 1.
 KW Oxidoreductase; Dioxygenase; Aromatic hydrocarbons catabolism; Iron.
 FT METAL 146 146 IRON (BY SIMILARITY).
 FT METAL 210 210 IRON (BY SIMILARITY).
 FT METAL 260 260 IRON (BY SIMILARITY).
 SQ SEQUENCE 291 AA; 32081 MW; 104F189FE1EDDA6A CRC64;

Query Match 77.5%; Score 31; DB 1; Length 291;
 Best Local Similarity 80.0%; Pred. No. 1.2e+02;
 Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 DFGWG 5
 DB 260 DFGWG 264

RESULT 13

BPHC_PSES1
 ID BPHC_PSES1 STANDARD; PRT; 292 AA.
 AC P17297; O52441;
 DT 01-AUG-1990 (Rel. 15, Created)
 DT 01-NOV-1995 (Rel. 32, Last sequence update)
 DT 15-JUL-1999 (Rel. 38, Last annotation update)
 DE BIPIHENTL-2,3-DIOL 1,2-DIOXYGENASE (EC 1.13.11.39) (230HBP OXYGENASE)
 GN BPHC.
 OS Pseudomonas sp. (strain KKS102).
 CC Bacteria; Proteobacteria.
 CC NCBI_TaxID=307;
 OX NCBI_TaxID=307;
 RN [1]
 RP SEQUENCE FROM N.A.

RX MEDLINE=89213965; PubMed=2540155;
 RA Kimbara K., Hashimoto T., Fukuda M., Koana T., Takagi M., Oishi M., Yano K.;
 RT "Cloning and sequencing of two tandem genes involved in degradation of 2,3-dihydroxybiphenyl to benzoic acid in the polychlorinated biphenyl-degrading soil bacterium Pseudomonas sp. strain KKS102.";
 RL J. Bacteriol. 171:2740-2747(1989).

RN [2]
 RP X-RAY CRYSTALLOGRAPHY (1.8 ANGSTROMS).
 RX MEDLINE=96226036; PubMed=8636975;
 RA Senda T., Sugiyama K., Narita H., Yamamoto T., Kimbara K., Fukuda M., Sato M., Yano K., Mitsui Y.;

RT "Three-dimensional structures of free form and two substrate complexes of an extradiol ring-cleavage type dioxygenase, the BphC enzyme from Pseudomonas sp. strain KKS102.";
 RL J. Mol. Biol. 255:735-752(1996).

CC -I- CATALYTIC ACTIVITY: BIPIHENTL-2,3-DIOL + O(2) = 2-HYDROXY-6-OXO-6-PHENYLHEXA-2,4-DIENATE + H(2)O.

CC -I- COFACTOR: FERROUS ION.
 CC -I- PATHWAY: DEGRADATION OF BIPIHENYLS AND POLYCHLOROBIPHENYLS (PCB) TO BENZOIC ACID AND CHLOROBENZOIC ACIDS.

CC -1- SUBUNIT: HOMOOCTAMER.
CC -1- SIMILARITY: BELONGS TO THE EXTRADIOL RING-CLEAVAGE DIOXYGENASE
CC FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: M26433; AAA25750.1; -
CC EMBL: D17319; BAA04141.1; -
CC PIR: A32312; DAPSPC.
CC PDB: 1DHY; 15-OCT-95.
CC InterPro: IPR000486; Extradiol_dioxygenase.
CC InterPro: IPR000325; Glyoxalase_1.
CC Pfam: PF00903; Glyoxalase; 2.
CC ProDom: PD000977; Extradiol_dioxygenase; 1.
CC PROSITE: PS00082; EXTRADIOL_DIOXYGENAS; 1.
CC Oxidoreductase: Dioxigenase; Aromatic hydrocarbons catabolism; Iron;
KM 3D-structure.
FT INIT_MER 0 0
FT METAL 145 145 IRON.
FT METAL 209 209 IRON.
FT METAL 260 260 IRON.
SQ SEQUENCE 292 AA; 32113 MW; 5D7D912F79EA8476 CRC64;

Query Match 77.5%; Score 31; DB 1; Length 292;
Best Local Similarity 80.0%; Pred. No. 1.2e+02;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 DFGWG 5
:||||
Db 260 EFGWG 264

RESULT 14
BPHC_PSEPA
ID BPHC_PSEPA STANDARD; PRT; 299 AA.
AC P11122;
DT 01-JUL-1989 (Rel. 11, Created)
DT 01-JUL-1989 (Rel. 11, Last annotation update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE BIPIENYL-2,3-DIOL 1,2-DIOXYGENASE (EC 1.13.11.35) (23OHP OXYGENASE)
DE (2,3-DIHYDROXYBIPIENYL DIOXYGENASE) (DHBD).
GN BPHC.
OS Pseudomonas paucimobilis (Sphingomonas paucimobilis).
OC Bacteria; Proteobacteria; alpha subdivision; Sphingomonadaceae;
OC Sphingomonas.
OX NCBI_TaxID=13689;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-Q1;
RC MEDLINE=88326911; PubMed=3137968;
RA Taira K., Hayase N., Arimura N., Yamashita S., Miyazaki T.,
RA Furukawa K.;
RT Cloning and nucleotide sequence of the 2,3-dihydroxybiphenyl
RT dioxygenase gene from the PCB-degrading strain of Pseudomonas
RT paucimobilis O1.";
RL Biochemistry 27:3990-3996(1988).
CC -1- CATALYTIC ACTIVITY: BIPIENYL-2,3-DIOL + O(2) -> 2-HYDROXY-6-OXO-
CC 6-PHENYLHEXA-2,4-DIENATE + H(2)O.
CC -1- COFACTOR: FERROUS ION.
CC -1- PATHWAY: DEGRADATION OF BIPIENYLS AND POLYCHLOROBIPHENYLS (PCB) TO
CC BENZOIC ACID AND CHLOROBENZOIC ACIDS.
CC -1- SUBUNIT: HOMOOCTAMER.
CC -1- SIMILARITY: BELONGS TO THE EXTRADIOL RING-CLEAVAGE DIOXYGENASE
CC FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: M20640; AAA25678.1; -
CC PIR: A28718; A28718.
CC HSSP: P17297; 1DHY.
CC InterPro: IPR000486; Extradiol_dioxygenase.
CC InterPro: IPR000325; Glyoxalase_1.
CC Pfam: PF00903; Glyoxalase; 2.
CC ProDom: PD000977; Extradiol_dioxygenase; 1.
CC PROSITE: PS00082; EXTRADIOL_DIOXYGENAS; 1.
CC Oxidoreductase: Dioxigenase; Aromatic hydrocarbons catabolism; Iron.
KM METAL 149 149 IRON (BY SIMILARITY).
FT METAL 212 212 IRON (BY SIMILARITY).
FT METAL 263 263 IRON (BY SIMILARITY).
SQ SEQUENCE 299 AA; 33095 MW; 29746CAB79D6FC7B CRC64;

Query Match 77.5%; Score 31; DB 1; Length 299;
Best Local Similarity 80.0%; Pred. No. 1.3e+02;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 DFGWG 5
:||||
Db 263 EFGWG 267

RESULT 15
HEMX_ECOLI
ID HEMX_ECOLI STANDARD; PRT; 393 AA.
AC P09127;
DT 01-MAR-1989 (Rel. 10, Created)
DT 01-MAR-1989 (Rel. 10, Last sequence update)
DT 20-AUG-2001 (Rel. 40, Last annotation update)
DE PUTATIVE UROPORPHYRIN-III C-METHYLTRANSFERASE (EC 2.1.1.107) (UROGEN
DE III METHYLASE) (ORF X).
GN HEMX OR B3803.
OS Escherichia coli.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Escherichia.
OX NCBI_TaxID=562;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-K12;
RX MEDLINE=89098348; PubMed=3062586;
RA Sasarman A., Echeland Y., Letowski J., Tardif D., Drolet M.,
RT "Nucleotide sequence of the hemX gene, the third member of the Uro
RT operon of Escherichia coli K12";
RL Nucleic Acids Res. 16:11835-11835(1988).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-K12 / CS520;
RX MEDLINE=89041586; PubMed=3054815;
RA Alefounder P.R.;
RT "The sequence of hemC, hemD and two additional E. coli genes";
RL Nucleic Acids Res. 16:9871-9871(1988).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN-K12 / MG1655;
RX MEDLINE=92358234; PubMed=1379743;
RA Daniels D.L., Plunkett G. III, Burland V.D., Blattner F.R.;
RT "Analysis of the Escherichia coli genome: DNA sequence of the region
RT from 84.5 to 86.5 minutes";
RL Science 257:771-778(1992).
RN [4]
RP SEQUENCE OF 1-12.
RC STRAIN-K12 / EMG2;
RX MEDLINE=97443975; PubMed=9298646;
RA Link A.J., Robison K., Church G.M.;

```

RT      "Comparing the predicted and observed properties of proteins encoded
RL      in the genome of Escherichia coli K-12."
CC      Electrophoresis 18:1259-1313(1997).
CC      -1- CATALYTIC ACTIVITY: 2 S-ADENOSYL-L-METHIONINE + UROPORPHYRIN III
CC      = 2 S-ADENOSYL-L-HOMOCYSTEINE + SIROHYDROCHLORIN.
CC      -1- PATHWAY: INVOLVED IN THE BIOSYNTHESIS OF SIROHEME AND COBALAMIN.
CC      -----
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC      or send an email to license@isb-sib.ch).
CC      -----
DR      EMBL; X13406; CAA31772.1; -
DR      EMBL; X12614; CAA31134.1; -
DR      EMBL; M87049; AAA67599.1; -
DR      EMBL; AE000456; AAC76806.1; -
DR      PIR; S02185; S02185.
DR      PIR; S30693; S30693.
DR      ECO2DBASE; B043.0; 6TH EDITION.
DR      ECO2DBASE; B043.1; 6TH EDITION.
DR      EcoGene; EG10433; hemX.
KW      Porphyrin biosynthesis; Transferase; Methyltransferase;
KW      Complete proteome.
SQ      SEQUENCE 393 AA; 42963 MW; 9D272C6401D0E354 CRC64;

```

```

Query Match          77.5%; Score 31; DB 1; Length 393;
Best Local Similarity 80.0%; Pred No. 1.6e+02;
Matches 4; Conservative 1; Mismatches 0; Indels 0;

```

```

Oy      2 FGWGX 6
       :||||
DB      55 YGWGX 59

```

Search completed: November 5, 2001, 12:58:55
Job time: 116 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 5, 2001, 12:56:59 ; Search time 22.88 Seconds
(without alignments)
38.358 Million cell updates/sec

Title: US-08-894-356C-21
Perfect score: 40
Sequence: 1 DFGWGK 6

Scoring table: BIOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 473505 seqs, 146272329 residues

Total number of hits satisfying chosen parameters: 473505

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 08
Maximum Match 100%
Listing first 45 summaries

Database :
1: SP-archaea:.*
2: SP-bacteria:.*
3: SP-fungi:.*
4: SP-human:.*
5: SP-invertebrate:.*
6: SP-mammal:.*
7: SP-mhc:.*
8: SP-organella:.*
9: SP-phage:.*
10: SP-plant:.*
11: SP-rodent:.*
12: SP-virus:.*
13: SP-vertebrate:.*
14: SP-unclassified:.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	40	100.0	219	10	Q9M5K9 euphorbia e
2	40	100.0	434	10	Q9FH97 arabidopsis
3	40	100.0	435	10	Q9FYM0 arabidopsis
4	40	100.0	436	10	Q9FYM1 arabidopsis
5	40	100.0	439	10	Q9ZTK5 arabidopsis
6	40	100.0	442	10	Q9FFA3 arabidopsis
7	40	100.0	443	10	Q9SND9 arabidopsis
8	40	100.0	443	10	Q9FJNO arabidopsis
9	40	100.0	446	10	Q9MBC1 arabidopsis
10	40	100.0	448	10	Q9MBD4 arabidopsis
11	40	100.0	448	10	Q9FID1 arabidopsis
12	40	100.0	448	10	Q9FH98 arabidopsis
13	40	100.0	449	10	Q9IJR4 arabidopsis
14	40	100.0	450	10	Q9SNE0 arabidopsis
15	40	100.0	451	10	Q9SNE0 arabidopsis
16	40	100.0	455	10	Q9J094 cucumis mel
17	40	100.0	455	10	Q9J094 cucumis mel
18	40	100.0	458	10	Q9J094 cucumis mel
19	40	100.0	460	10	Q43583 nicotiana gl

20	40	100.0	461	10	Q9LR83 arabidopsis
21	40	100.0	469	10	Q9ZWR8 gentiana tr
22	40	100.0	475	10	Q9LFB5 arabidopsis
23	40	100.0	479	10	Q9MBD5 arabidopsis
24	40	100.0	553	10	Q9FVY9 gentiana tr
25	37	92.5	99	10	Q23943 gentiana tr
26	37	92.5	180	10	Q9LR06 arabidopsis
27	37	92.5	293	10	Q9FSP0 arabidopsis
28	37	92.5	399	2	Q9I458 pseudomonas
29	37	92.5	431	10	Q9SST8 ipomoea bat
30	37	92.5	433	10	Q9FI78 arabidopsis
31	37	92.5	443	10	Q9FI40 arabidopsis
32	37	92.5	451	10	Q9LR08 arabidopsis
33	37	92.5	451	10	Q9LR07 arabidopsis
34	37	92.5	452	10	Q9FVFI fragaria an
35	37	92.5	452	10	Q9FNP9 arabidopsis
36	37	92.5	455	10	Q9FID3 arabidopsis
37	37	92.5	463	10	Q9FID2 arabidopsis
38	37	92.5	482	10	Q04201 arabidopsis
39	36	90.0	280	10	Q9I365 prunus arme
40	36	90.0	426	10	Q9FFQ7 arabidopsis
41	36	90.0	430	10	Q9SMW7 arabidopsis
42	36	90.0	441	10	Q9FLM5 arabidopsis
43	35	87.5	363	2	Q9A0Q7 streptococ
44	35	87.5	428	10	Q9FLW4 arabidopsis
45	35	87.5	430	10	Q9SQ02 clarkia con

ALIGNMENTS

RESULT 1
ID Q9M5K9 PRELIMINARY: PRT: 219 AA.
AC Q9M5K9:
DT 01-OCT-2000 (TREMBLrel. 15, Created)
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
DE F21J9_20-LIKE PROTEIN (FRAGMENT).
OS Euphorbia esula (leafy spurge).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
OC eurosids I; Malpighiales; Euphorbiaceae; Euphorbia.
OX NCBI_TaxID=3993;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=UNDERGROUND ADVENTITIOUS BUDS;
RA Anderson J.V., Horvath D.P.;
RT "Identification of mRNAs expressed in underground adventitious buds of
RT Euphorbia esula (leafy spurge)."
RL Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF227981; AAF34801.1;
FT NON_TER 1
SQ SEQUENCE 219 AA: 24236 MW: DDD96C93FC87BC30 CRC64:

Query Match 100.0%; Score 40; DB 10; Length 219;
Best local Similarity 100.0%; Pred. No. 8.5;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DFGWGK 6
DB 155 DFGWGK 161

RESULT 2
ID Q9FH97 PRELIMINARY: PRT: 434 AA.
AC Q9FH97:
DT 01-MAR-2001 (TREMBLrel. 16, Created)
DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)
DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
DE ANTHANILATE N-HYDROXYCINNAOYL/BENZOYLTRANSFERASE-LIKE PROTEIN.

OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC eustosids II; Brassicales; Brassicaceae; Arabidopsis.
 OK NCBI_TaxID=3702;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=COLUMBIA;
 RX MEDLINE=20181125; PubMed=10718197;
 RA Sato S., Nakamura Y., Kaneko T., Katoh T., Asamizu E., Kotani H.,
 RA Tabata S.;
 RT "Structural analysis of Arabidopsis thaliana chromosome 5. X. Sequence
 RT features of the regions of 3,076,755 bp covered by sixty P1 and TAC
 RT clones";
 RL DNA Res. 7:31-63(2000).
 DR EMBL; AB020742; BAB10950.1; -
 DR InterPro: IPR003480; Transferase.
 DR Pfam: PF02458; Transferase; 1.
 KW Transferase
 SO SEQUENCE 434 AA; 48635 MW; 7E4C4C61DA1624AA CRC64;

Query Match 100.0%; Score 40; DB 10; Length 434;
 Best Local Similarity 100.0%; Pred. No. 18;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 DEGCK 6
 |||||

DB 376 DEGCK 381

RESULT 3

O9FYMO PRELIMINARY; PRT; 435 AA.
 AC O9FYMO:
 DT 01-MAR-2001 (TREMBLrel. 16, Created)
 DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)
 DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
 DE F21J9.9.
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC eustosids II; Brassicales; Brassicaceae; Arabidopsis.
 OK NCBI_TaxID=3702;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Johnson-Hopson C., Brooks S., Buehler E., Chao Q., Khan S., Kim C.,
 RA Shinn P., Altafi H., Bel O., Chin C., Chlou J., Choi E., Conn L.,
 RA Conway A., Gonzales A., Hansen N., Howling B., Koo T., Lam B., Lee J.,
 RA Lenz C., Li J., Liu A., Liu K., Liu S., Mukharzky N., Nguyen M.,
 RA Palm C., Pham P., Sakano H., Schwartz J., Southwick A., Thayerl A.,
 RA Toriumi M., Vaysberg M., Yu G., Federspiel N.A., Theologis A.,
 RA Ecker J.R.;
 RT "Genomic sequence for Arabidopsis thaliana BAC F21J9 from chromosome
 RT 1";
 RL Submitted (DEC-1996) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RA Ecker J.R.;
 RL Submitted (DEC-1996) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A.
 RA Ecker J.R.;
 RL Submitted (FEB-1997) to the EMBL/GenBank/DBJ databases.
 RN [4]
 RP SEQUENCE FROM N.A.
 RA Ecker J.R.;
 RL Submitted (MAR-1997) to the EMBL/GenBank/DBJ databases.
 RN [5]
 RP SEQUENCE FROM N.A.
 RA Ecker J.R.;
 RL Submitted (JUN-1997) to the EMBL/GenBank/DBJ databases.

RP SEQUENCE FROM N.A.
 RA Ecker J.R.;
 RL Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.
 RN [7]
 RP SEQUENCE FROM N.A.
 RA Ecker J.R.;
 RL Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.
 RN [8]
 RP SEQUENCE FROM N.A.
 RA Cheuk R., Shinn P., Brooks S., Buehler E., Chao Q., Johnson-Hopson C.,
 RA Khan S., Kim C., Altafi H., Bel O., Chin C., Chlou J., Choi E., Conn L.,
 RA Conway A., Gonzales A., Hansen N., Howling B., Koo T., Lam B.,
 RA Lee J., Lenz C., Li J., Liu A., Liu K., Liu S., Mukharzky N.,
 RA Nguyen M., Palm C., Pham P., Sakano H., Schwartz J., Southwick A.,
 RA Thayerl A., Toriumi M., Vaysberg M., Yu G., Davis R., Federspiel N.,
 RA Theologis A., Ecker J.;
 RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AC000103; AAF97979.1; -
 DR InterPro: IPR003480; Transferase.
 DR Pfam: PF02458; Transferase; 1.
 SO SEQUENCE 435 AA; 49288 MW; B369CEC11D982DF CRC64;

Query Match 100.0%; Score 40; DB 10; Length 435;
 Best Local Similarity 100.0%; Pred. No. 18;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 DEGCK 6
 |||||

DB 369 DEGCK 374

RESULT 4

O9FYML PRELIMINARY; PRT; 436 AA.
 AC O9FYML:
 DT 01-MAR-2001 (TREMBLrel. 16, Created)
 DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)
 DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
 DE F21J9.8.
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC eustosids II; Brassicales; Brassicaceae; Arabidopsis.
 OK NCBI_TaxID=3702;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Johnson-Hopson C., Brooks S., Buehler E., Chao Q., Khan S., Kim C.,
 RA Shinn P., Altafi H., Bel O., Chin C., Chlou J., Choi E., Conn L.,
 RA Conway A., Gonzales A., Hansen N., Howling B., Koo T., Lam B., Lee J.,
 RA Lenz C., Li J., Liu A., Liu K., Liu S., Mukharzky N., Nguyen M.,
 RA Palm C., Pham P., Sakano H., Schwartz J., Southwick A., Thayerl A.,
 RA Toriumi M., Vaysberg M., Yu G., Federspiel N.A., Theologis A.,
 RA Ecker J.R.;
 RT "Genomic sequence for Arabidopsis thaliana BAC F21J9 from chromosome
 RT 1";
 RL Submitted (DEC-1996) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RA Ecker J.R.;
 RL Submitted (DEC-1996) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A.
 RA Ecker J.R.;
 RL Submitted (FEB-1997) to the EMBL/GenBank/DBJ databases.
 RN [4]
 RP SEQUENCE FROM N.A.
 RA Ecker J.R.;
 RL Submitted (MAR-1997) to the EMBL/GenBank/DBJ databases.
 RN [5]
 RP SEQUENCE FROM N.A.
 RA Ecker J.R.;
 RL Submitted (JUN-1997) to the EMBL/GenBank/DBJ databases.

RN [6]
RP SEQUENCE FROM N.A.
RA Ecker J.R.;
RN Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.
RL [7]
RP SEQUENCE FROM N.A.
RA Ecker J.R.;
RN Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.
RL [8]
RP SEQUENCE FROM N.A.
RA Chen R., Shinn P., Brooks S., Buehler E., Chao Q., Johnson-Hopson C.,
RA Khan S., Kim C., Altaji H., Bel B., Chin C., Chou J., Choi E.,
RA Com L., Conway A., Gonzalez A., Hansen N., Howing B., Koo T., Lam B.,
RA Lee J., Lenz C., Li J., Liu A., Liu J., Liu S., Mukharsky N.,
RA Nguyen M., Palm C., Pham P., Sakano H., Schwartz J., Southwick A.,
RA Theverl A., Toriumi M., Vaysberg M., Yu G., Davis R., Federspiel N.,
RN Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.
RL EMBL: AC000103; AAF97955.1; -;
DR InterPro: IPR003480; Transferase.
DR Pfam: PF02458; Transferase; 1.
SO SEQUENCE 436 AA; 48824 MW; B303A52DE3FD1355 CRC64;

Query Match 100.0%; Score 40; DB 10; Length 436;
Best Local Similarity 100.0%; Pred. No. 18;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 DFGWGK 6
Db 372 DFGWGK 377

RESULT 5
ID 092TK5 PRELIMINARY; PRT; 439 AA.
AC 092TK5;
DT 01-MAY-1999 (TREMblrel. 10, Created)
DT 01-MAY-1999 (TREMblrel. 10, Last sequence update)
DT 01-JUN-2001 (TREMblrel. 17, Last annotation update)
DE DEACTYLAVINDOLINE 4-O-ACETYLTRANSFERASE (EC 2.3.1.107).
GN DAT.
OS Catharanthus roseus (Rosy periwinkle) (Madagascar periwinkle).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC Asteridae; eusterids I; Gentianales; Apocynaceae; Catharanthus.
OX NCBI_TaxID=4058;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-CV LITTLE DELICATA;
RX MEDLINE-98346012; PubMed-9681034;
RA St-Pierre B., Laflamme P., Alarco A.M., De Luca V.;
RT "The terminal O-acetyltransferase involved in vindoline biosynthesis
RT defines a new class of proteins responsible for coenzyme A-dependent
RT acyl transfer.";
RL Plant J. 14:703-713(1998).
DR EMBL: AF053307; AAC9311.1; -;
DR InterPro: IPR003480; Transferase.
DR Pfam: PF02458; Transferase; 1.
KW Transferase; Acyltransferase.
SO SEQUENCE 439 AA; 49873 MW; 1357C76550E0D04D CRC64;

Query Match 100.0%; Score 40; DB 10; Length 439;
Best Local Similarity 100.0%; Pred. No. 18;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 DFGWGK 6
Db 380 DFGWGK 385

RESULT 6

O9FF43
ID O9FF43 PRELIMINARY; PRT; 442 AA.
AC O9FF43;
DT 01-MAR-2001 (TREMblrel. 16, Created)
DT 01-MAR-2001 (TREMblrel. 16, Last sequence update)
DT 01-JUN-2001 (TREMblrel. 17, Last annotation update)
DE ANTHRANILATE N-HYDROXYCINNAMOYL/BENZOYLTRANSFERASE-LIKE PROTEIN.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eustoids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-COLUMBIA;
RX MEDLINE-97471969; PubMed-9330910;
RA Sato S., Kotani H., Nakamura Y., Kaneko T., Asamizu E., Fukami M.,
RA Miyajima N., Tabata S.;
RT "Structural analysis of Arabidopsis thaliana chromosome 5. I. Sequence
RT features of the 1.6 Mb regions covered by twenty physically assigned
RT P1 clones.";
RL DNA Res. 4:215-230(1997).
DR EMBL: AB005247; BAB11280.1; -;
DR InterPro: IPR003480; Transferase.
DR Pfam: PF02458; Transferase; 1.
KW Transferase.
SO SEQUENCE 442 AA; 49491 MW; 1D280C885875DA47 CRC64;

Query Match 100.0%; Score 40; DB 10; Length 442;
Best Local Similarity 100.0%; Pred. No. 18;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 DFGWGK 6
Db 384 DFGWGK 389

RESULT 7
ID O9SND9 PRELIMINARY; PRT; 443 AA.
AC O9SND9;
DT 01-MAY-2000 (TREMblrel. 13, Created)
DT 01-MAY-2000 (TREMblrel. 13, Last sequence update)
DT 01-JUN-2001 (TREMblrel. 17, Last annotation update)
DE ANTHRANILATE N-HYDROXYCINNAMOYL/BENZOYLTRANSFERASE-LIKE PROTEIN.
GN P130.120.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eustoids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RA Barques M., Collado M.C., Navarro P., Terol J., Perez-Alonso M.,
RA Mewes H.W., Lemcke K., Mayer K.F.X., Queller F., Salanoubat M.;
RT Submitted (NOV-1999) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC EU Arabidopsis sequencing project;
RL Submitted (DEC-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL: AL132976; CAB62307.1; -;
DR InterPro: IPR003480; Transferase.
DR Pfam: PF02458; Transferase; 1.
KW Transferase.
SO SEQUENCE 443 AA; 49960 MW; C4ADB67A0E3AB3D4 CRC64;

Query Match 100.0%; Score 40; DB 10; Length 443;
Best Local Similarity 100.0%; Pred. No. 18;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 DFGWGK 6

Db 386 DFGWK 391

RESULT 8

ID Q9FJNO PRELIMINARY; PRT: 443 AA.

AC Q9FJNO; 01-MAR-2001 (TREMBlrel. 16, Created)
 DT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)
 DE 01-JUN-2001 (TREMBlrel. 17, Last annotation update)
 DE N-HYDROXYCINNAMOYL/BENZOYLTRANSFERASE.
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC eurosids II; Brassicales; Brassicaceae; Arabidopsids.
 GN NCBI_TaxID=3702;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=COLUMBIA;
 RX MEDLINE=98403884; PubMed=9734815;
 RA Kotani H., Nakamura Y., Sato S., Asanizu E., Kaneko T., Miyajima N.,
 RA Tabata S.;
 RT "Structural analysis of Arabidopsis thaliana chromosome 5. VI.
 RT Sequence features of the regions of 1,367,185 bp covered by 19
 RT physically assigned P1 and TAC clones."
 RL DNA Res. 5:203-216(1998).
 DR EMBL: AB013396; BAB08854.1; -
 DR InterPro: IPR003480; Transferase.
 DR Pfam: PF02458; Transferase; 1.
 KW Transferase.
 SO SEQUENCE 443 AA; 49950 MW; 16FFC01B99035A5D CRC64;

Query Match 100.0%; Score 40; DB 10; Length 443;
 Best Local Similarity 100.0%; Pred. No. 18;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 DFGWK 6
 Db 385 DFGWK 390

RESULT 9
 ID Q9MBC1 PRELIMINARY; PRT: 446 AA.

AC Q9MBC1; 01-OCT-2000 (TREMBlrel. 15, Created)
 DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
 DE 01-JUN-2001 (TREMBlrel. 17, Last annotation update)
 DE ANTHOCYANIN ACYLTRANSFERASE (FRAGMENT).
 OS Perilla frutescens.
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 OC Asteridae; euasterids I; Lamiales; Lamiaceae; Perilla.
 GN NCBI_TaxID=48386;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=LEAF;
 RA Sakakibara K.Y., Tanaka Y., Mizutani M.F., Fujiwara H., Fukui Y.,
 RA Ashikari T., Yamaguchi M., Kusumi T.;
 RT "Molecular and biochemical characterization of a novel
 RT hydroxycinnamoyl-CoA: anthocyanin 3-O-glucoside-6-";
 RT Submitted (JUN-1999) to the EMBL/Genbank/DBJ databases.
 RL EMBL: AB029340; BAA93475.1; -
 DR InterPro: IPR003480; Transferase.
 DR Pfam: PF02458; Transferase; 1.
 KW Transferase; Acyltransferase.
 FT NON_TER 1
 SO SEQUENCE 446 AA; 50675 MW; DC83C69AF248A593 CRC64;

Query Match 100.0%; Score 40; DB 10; Length 446;

Best Local Similarity 100.0%; Pred. No. 19;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 DFGWK 6
 Db 389 DFGWK 394

RESULT 10

ID Q9MBD4 PRELIMINARY; PRT: 448 AA.

AC Q9MBD4; 01-OCT-2000 (TREMBlrel. 15, Created)
 DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
 DE 01-JUN-2001 (TREMBlrel. 17, Last annotation update)
 DE ACYLTRANSFERASE HOMOLOG.
 GN PAT48.
 OS Petunia hybrida (Petunia).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 OC Asteridae; euasterids I; Solanales; Solanaceae; Petunia.
 GN NCBI_TaxID=4102;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Yonekura-Sakakibara K., Tanaka Y., Fukuchi-Mizutani M., Fujiwara H.,
 RA Fukui Y., Toshihiko A., Yamaguchi M., Kusumi T.;
 RT "Molecular cloning and biochemical characterization of
 RT hydroxycinnamoyl-CoA:anthocyanin 3-O-glucoside-6-O-
 RT hydroxycinnamoyltransferase from Perilla frutescens and diverse plant
 RT acyltransferase homologs."
 RL Submitted (APR-1999) to the EMBL/Genbank/DBJ databases.
 DR EMBL: AB026495; BAA93451.1; -
 DR InterPro: IPR003480; Transferase.
 DR Pfam: PF02458; Transferase; 1.
 KW Transferase.
 SO SEQUENCE 448 AA; 49919 MW; 85D8EECEDDC188C CRC64;

Query Match 100.0%; Score 40; DB 10; Length 448;
 Best Local Similarity 100.0%; Pred. No. 19;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 DFGWK 6
 Db 390 DFGWK 395

RESULT 11

ID Q9FTD1 PRELIMINARY; PRT: 448 AA.

AC Q9FTD1; 01-MAR-2001 (TREMBlrel. 16, Created)
 DT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)
 DE 01-JUN-2001 (TREMBlrel. 17, Last annotation update)
 DE ANTHOCYANIN ACYLTRANSFERASE-LIKE PROTEIN.
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 OC eurosids II; Brassicales; Brassicaceae; Arabidopsids.
 GN NCBI_TaxID=3702;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=COLUMBIA;
 RX MEDLINE=99156233; PubMed=10048488;
 RA Asanizu E., Sato S., Kaneko T., Nakamura Y., Kotani H., Miyajima N.,
 RA Tabata S.;
 RT "Structural analysis of Arabidopsis thaliana chromosome 5. VIII.
 RT Sequence features of the regions of 1,081,958 bp covered by seventeen
 RT physically assigned P1 and TAC clones."
 RL DNA Res. 5:379-391(1998).
 DR EMBL: AB016892; BAB10831.1; -
 DR InterPro: IPR003480; Transferase.
 DR Pfam: PF02458; Transferase; 1.

KW Transferrase; Acyltransferase.
SQ SEQUENCE 448 AA; 49916 MW; 27878BBE9B8984E CRC64;

Query Match 100.0%; Score 40; DB 10; Length 448;
Best Local Similarity 100.0%; Pred. No. 19;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 DFGWGK 6
DB 393 DFGWGK 398

RESULT 12

ID 09FH98 PRELIMINARY; PRT; 448 AA.
AC 09FH98;
DT 01-MAR-2001 (TREMBlrel. 16, Created)
DT 01-MAR-2001 (TREMBlrel. 17, Last annotation update)
DE ANTHRANILATE N-HYDROXYCINNAMOYL/BENZOYLTRANSFERASE-LIKE PROTEIN.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC euroids II; Brassicales; Brassicaceae; Arabidopsis.
NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-COLUMBIA;
RX MEDLINE-20181125; PubMed-10718197;
RA Sato S., Nakamura Y., Kaneko T., Katoh T., Asamizu E., Kotani H.,
RA Tabata S.;
RT "Structural analysis of Arabidopsis thaliana chromosome 5. X. Sequence
RT features of the regions of 3,076,755 bp covered by sixty P1 and TAC
RT clones.";
RL DNA Res. 7:31-63(2000).
DR EMBL; AB020742; BAB10949.1; -;
DR InterPro; IPR003480; Transferase.
DR Pfam; PF02458; Transferase; 1.
KW Transferrase.
SQ SEQUENCE 448 AA; 49939 MW; 04EB676545093DEA CRC64;

Query Match 100.0%; Score 40; DB 10; Length 448;
Best Local Similarity 100.0%; Pred. No. 19;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 DFGWGK 6
DB 389 DFGWGK 394

RESULT 13

ID 09LJB4 PRELIMINARY; PRT; 449 AA.
AC 09LJB4;
DT 01-OCT-2000 (TREMBlrel. 15, Created)
DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
DT 01-JUN-2001 (TREMBlrel. 17, Last annotation update)
DE ANTHOCYANIN 5-AROMATIC ACYLTRANSFERASE-LIKE PROTEIN.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC euroids II; Brassicales; Brassicaceae; Arabidopsis.
NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-COLUMBIA;
RA Kaneko T., Kato T., Sato S., Nakamura Y., Asamizu E., Tabata S.;
RL Submitted (OCT-1999) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-COLUMBIA;

PUBMED-10907853;

RA Nakamura Y.;

RT "Structural analysis of Arabidopsis thaliana chromosome 3. II.

RT Sequence features of the regions of 4,251,695 bp covered by ninety P1,

RT TAC and BAC clones.";

RL DNA Res. 7:217-221(2000).

DR EMBL; AP000606; BAB01191.1; -;

DR InterPro; IPR003480; Transferase.

DR Pfam; PF02458; Transferase; 1.

KW Transferrase; Acyltransferase.
SQ SEQUENCE 449 AA; 49439 MW; 2446D502AEC9D99B CRC64;

Query Match 100.0%; Score 40; DB 10; Length 449;
Best Local Similarity 100.0%; Pred. No. 19;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 DFGWGK 6
DB 394 DFGWGK 399

RESULT 14

ID 09SNE0 PRELIMINARY; PRT; 450 AA.
AC 09SNE0;
DT 01-MAY-2000 (TREMBlrel. 13, Created)
DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
DT 01-JUN-2001 (TREMBlrel. 17, Last annotation update)
DE ANTHRANILATE N-HYDROXYCINNAMOYL/BENZOYLTRANSFERASE-LIKE PROTEIN.
GN F1C1.110.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC euroids II; Brassicales; Brassicaceae; Arabidopsis.
NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RA Barques M., Collado M.C., Navarro P., Terol J., Perez-Alonso M.,
RA Mewes H.W., Lemcke K., Mayer K.F.X., Queitner F., Salanoubat M.;
RL Submitted (NOV-1999) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA EU Arabidopsis sequencing project;
RL Submitted (DEC-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AL132976; CAB62306.1; -;
DR InterPro; IPR003480; Transferase.
DR Pfam; PF02458; Transferase; 1.
KW Transferrase.
SQ SEQUENCE 450 AA; 49805 MW; 5DA4AAAABCD3F0C6 CRC64;

Query Match 100.0%; Score 40; DB 10; Length 450;
Best Local Similarity 100.0%; Pred. No. 19;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 DFGWGK 6
DB 391 DFGWGK 396

RESULT 15

ID 064470 PRELIMINARY; PRT; 451 AA.
AC 064470;
DT 01-AUG-1998 (TREMBlrel. 07, Created)
DT 01-AUG-1998 (TREMBlrel. 07, Last sequence update)
DT 01-JUN-2001 (TREMBlrel. 17, Last annotation update)
DE PUTATIVE ANTHRANILATE N-HYDROXYCINNAMOYL/BENZOYLTRANSFERASE.
GN T20K24.8.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;

OC eurosids II: Brassicales; Brassicaceae; Arabidopsis.
 OX NCBI_TaxID-3702;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-CV. COLUMBIA;
 RA Rounsley S.D., Kaul S., Lin X., Ketchum K.A., Crosby M.L.,
 RA Brandon R.C., Sykes S.M., Mason T.M., Kerlavage A.R., Adams M.D.,
 RA Somerville C.R., Venter J.C.,
 RL Submitted (JUN-1998) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AC002392; AAD12025.1; -
 DR InterPro; IPR003480; Transferase.
 DR Pfam; PF02458; Transferase; 1.
 KW Transferase
 SO SEQUENCE 451 AA; 50434 MW; B13B2CBF4DB4482 CRC64;

Query Match 100.0%; Score 40; DB 10; Length 451;
 Best Local Similarity 100.0%; Pred. No. 19;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 DFGWCK 6
 |||||
 Db 397 DFGWCK 402

Search completed: November 5, 2001, 12:58:39
 Job time: 100 sec